



Espace de conception et modèle d'interaction multi-tactile gestuel : un environnement de développement pour enrichir le modèle

Rudy Morin

► To cite this version:

Rudy Morin. Espace de conception et modèle d'interaction multi-tactile gestuel : un environnement de développement pour enrichir le modèle. Autre [cs.OH]. Université Toulouse le Mirail - Toulouse II, 2011. Français. NNT : 2011TOU20025 . tel-00634026

HAL Id: tel-00634026

<https://theses.hal.science/tel-00634026>

Submitted on 20 Oct 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université
de Toulouse

THÈSE

En vue de l'obtention du DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par

Université Toulouse 2 Le Mirail (UT2 Le Mirail)

Discipline ou spécialité

Études audiovisuelles

Présentée et soutenue par

Rudy MORIN

le vendredi 24 juin 2011

Titre

Espace de conception et modèle d'interaction multi-tactile gestuel

Un environnement de développement pour enrichir le modèle

Ecole doctorale

Arts, Lettres, langues, Philosophie, Communication (ALLPH@)

Unité de recherche

LARA EA 4154

Directeur(s) de Thèse

Gilles MÉTHEL, Professeur, Université Toulouse II Le Mirail

Jean-Pierre JESSEL, Professeur, Université Toulouse III Paul Sabatier

Rapporteurs

Edmond COUCHOT, Professeur émérite, Université Paris VIII

Pascal GUITTON, Professeur, Université Bordeaux I

THÈSE

En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par
Université Toulouse 2 Le Mirail (UT2 Le Mirail)

Discipline ou spécialité
Études audiovisuelles

Présentée et soutenue par

Rudy MORIN

le vendredi 24 juin 2011

Titre

Espace de conception et modèle d'interaction multi-tactile gestuel
Un environnement de développement pour enrichir le modèle

Title

Design space and multi-touch gestural interaction model
A framework to enhance the model

École doctorale

Arts, Lettres, Langues, Philosophie, Communication (ALLPH@)

Unité de recherche

LARA EA 4154

Directeurs de Thèse

Gilles MÉTHEL, Professeur, Université Toulouse II Le Mirail
Jean-Pierre JESSEL, Professeur, Université Toulouse III Paul Sabatier

Rapporteurs

Edmond COUCHOT, Professeur émérite, Université Paris VIII
Pascal GUITTON, Professeur, Université Bordeaux I

Résumé

L'affinage technique et l'adoption récente des technologies tactiles multi-points par les industriels et les utilisateurs ont fixé l'attention des designers d'interaction sur ces technologies. Tandis que de nombreuses études en interaction homme-machine se sont intéressées à comparer la performance de ces interfaces à celle des interfaces WIMP traditionnelles, peu se sont attachées à intégrer dans leur approche les spécificités du canal gestuel et les modalités d'interactions multi-tactiles. Dans cette étude, je défends l'idée que le design de telles interactions ne peut être approché qu'en suivant un modèle d'interaction spécifique intégrant l'ensemble des composantes physiques, cognitives, sensorielles et motrices du geste dans le couplage homme-machine. J'articule ma recherche autour d'un espace de conception, courte analyse sociotechnique de mon objet d'étude, dans lequel je définis un modèle d'interaction descriptif et génératif. Je détermine un ensemble de principes conceptuels et techniques permettant l'évaluation et la conception du design d'interfaces multi-tactiles de manière systémique et extensible. Au cours de cette étude, je précise les limites du paradigme d'« interface naturelle » en nuancant les effets du réalisme des interactions dans l'efficacité de tels systèmes. Enfin, je présente les travaux de conception et de développement d'un environnement de développement réalisé dans le cadre d'un dispositif CIFRE qui a accompagné cette étude et permis d'enrichir le modèle théorique.

Mots-clefs : multi-tactile, design d'interaction, modèle d'interaction, interface homme-machine, interface utilisateur naturelle, NUI, geste, canal gestuel, hypperréalité, visibilité, affordance, retours utilisateur, environnement de développement, logiciel tiers, interface de développement, API, CIFRE.

Abstract

The technical refinement and the recent adoption of multi-touch technologies by both the industry and users made these technologies a major concern for interaction designers. While many studies on human-computer interaction have started to compare this type of interface with traditional WIMP interfaces performance-wise, few have included in their approach the gestural and system mapping specificities of multi-touch interactions. In this study, I defend the idea that such interaction design should follow a specific interaction model taking into account physical, cognitive, sensitive and motor aspects of gestures in the human-computer relation. I define a design space, sort of socio-technical approach, which participates in the definition of a descriptive and generative interaction model. I establish a set of conceptual and technical principles that allow the evaluation and conception of multi-touch interfaces, in a systemic and extensible way. Throughout this study, I emphasize the limits of the « natural user interface » paradigm by having a better understanding of how interaction realism affects system efficiency. Finally, I present a multi-touch framework developed as part of a CIFRE collaboration, which helped this study in the design and the extension of the conceptual model.

Keywords : multi-touch, interaction design, interaction model, human-computer interface, natural user interface, NUI, gesture, gestural channel, hyperrealism, visibility, affordance, user feedback, framework, middleware, application programming interface, API, CIFRE.

Laboratoire de recherche

LARA - Laboratoire de Recherches en Audiovisuel [E.A. 4154]
5 Allées Antonio Machado - 31058 Toulouse Cedex

REMERCIEMENTS

A Monsieur le Professeur Gilles Méthel

Recevez, Monsieur Méthel, mes plus sincères remerciements pour avoir accepté de diriger ma thèse. Je tiens à vous exprimer toute ma gratitude pour votre soutien sans réserve et la confiance que vous m'avez témoignée tout au long de ces années. Je me souviens encore du moment où nous nous sommes rencontrés, il y a de cela déjà 10 ans, quelque part à Castres, en nous demandant tous deux, d'un air circonspect, si la bonne fortune, pour chacun de nous, allait accompagner nos projets et nos aventures. Aujourd'hui, je peux vous garantir que la chance n'y a été pour rien. Seule votre ténacité, vos convictions idéologiques et votre engagement personnel ont permis, à mes compagnons de fortune et moi-même de faire si bon chemin. Je suis également profondément reconnaissant pour votre disponibilité, vos conseils avisés et les échanges enrichissants que nous avons pu avoir tout ce temps et que nous aurons encore, je l'espère. Soyez assuré, Monsieur, de toute mon estime et de mon profond respect.

A Monsieur le Professeur Jean-Pierre Jessel

Je vous adresse mes remerciements pour avoir accepté la codirection de ma thèse et pour l'intérêt que vous portez à mes travaux de recherche. Je vous exprime également toute ma reconnaissance pour votre investissement et vos précieuses recommandations sur la méthodologie et la rigueur scientifique qui se sont révélées nécessaires à la bonne conduite de mon projet de recherche. Je suis également profondément réjoui de vous avoir rencontré suffisamment tôt et d'avoir pu profiter de nos échanges sur la réalité virtuelle et les interfaces utilisateur émergentes, domaines de recherche qui nous passionnent tant. Merci Monsieur Jessel d'avoir partagé votre savoir avec tant de générosité.

A Monsieur le professeur Edmond Couchot

Je vous présente mes plus vifs remerciements pour l'intérêt que vous avez manifesté envers mon travail de recherche en acceptant d'être pré-rapporteur et membre du jury. Je suis réellement honoré de vous savoir présent lors de la soutenance. Veuillez recevoir, monsieur Couchot, l'expression de ma profonde estime.

A Monsieur le professeur Pascal Guitton

Je vous remercie vivement pour l'intérêt que vous avez accordé à mes travaux de recherche en acceptant d'être pré-rapporteur et membre du jury. Je suis réellement honoré de vous savoir présent lors de la soutenance. Veuillez recevoir, monsieur Guitton, l'expression de mon profond respect.

A Monsieur Jean-François Pal et Madame Cécile Dos Santos

Je vous remercie tous deux, Monsieur Pal et Madame Dos Santos, directeurs de l'entreprise Virtual-IT, de m'avoir accordé votre confiance et de l'intérêt que vous avez porté très tôt à mes travaux de recherche. Je suis profondément reconnaissant pour les moyens humains et financiers que vous n'avez pas hésité à mettre en œuvre afin de voir aboutir ces trois années de recherche et développement. Merci également pour cette formidable aventure humaine que nous avons partagé au sein de votre entreprise, de jour comme de nuit, dans la joie comme dans les moments plus difficiles. Je pense également à tous mes collègues de travail : Julien Agullo, Philippe Palacio, Philippe Henry, Simon Moisy et tous ceux qui m'ont réservé un accueil remarquable et qui ont partagé avec moi les joies et les affres de la recherche et du développement. Je remercie tout particulièrement Yohan Gregoire, Philippe Ercolessi et Eddy Chung qui m'ont apporté une aide précieuse en vérifiant la cohérence des parties techniques de cette étude.

A Monsieur Philippe Coste, directeur d'Epitech Toulouse

Je vous suis réellement reconnaissant, Monsieur Coste, de nous avoir fait confiance, alors que notre projet commençait tout juste à s'élaborer, en acceptant un partenariat entre Virtual-IT et votre école. J'espère que cette collaboration vous a été autant agréable et profitable qu'elle l'a été pour nous. Nous tenons à remercier tout particulièrement Monsieur Gildas Vinson, directeur pédagogique de l'école, qui a fourni un travail de suivi remarquable avec les groupes d'étudiants impliqués dans le partenariat. Nous sommes réellement reconnaissants du temps et des efforts que vous avez mobilisés et nous espérons que ce rapprochement avec Virtual-IT se poursuive avec la même dynamique.

Pour l'anecdote, je n'oublierai jamais ce qu'un jour vous m'avez dit, devant le projet LUNAR, lorsque nous le présentions sur un salon toulousain : « *Grâce à vous, je vis un rêve d'enfant.* » Car je crois sincèrement que c'est un sentiment que tout chercheur partage, en

son âme, le besoin motivé de suivre les chemins merveilleux d'un rêve d'enfant. Merci de m'avoir rappelé combien la recherche est avant tout une aventure extraordinairement humaine.

A mes parents, Gérard et Marguerite Morin

Je vous serai éternellement reconnaissant pour la confiance et le soutien sans faille que vous m'accordez depuis toujours. Merci pour ces instants en famille qui m'ont redonné courage et qui m'ont rappelé combien la simplicité et la beauté de ces moments étaient uniques. Vous rougissez tous deux de ne pas avoir eu la chance de faire de hautes études et pourtant, si vous saviez combien votre caractère et vos paroles reflètent la richesse de votre histoire, vous comprendriez à quel point je vous suis redevable de ce que vous m'avez transmis. Je vous dois bien plus que ce chemin jusqu'à la thèse, je vous dois l'intérêt ambivalent que je porte à l'art comme à la science, à la littérature comme au bricolage¹, au sensible comme au sensé. Merci pour votre obstination à croire en moi, merci pour tout.

A ma compagne Marine Coquelin

Merci à toi de m'avoir supporté pendant cette épreuve difficile de recherche et d'écriture, durant ces trois années de thèse, où notre temps passé ensemble était réglé par les contraintes professionnelles et parfois entaché par la fatigue liée à mes nombreux déplacements entre Toulouse, Grenoble, Paris, Tarbes, Montauban, Lavaur... Merci pour tous ces moments d'évasion à deux qui m'ont permis de prendre du recul sur mon travail et qui m'ont donné le courage d'aller jusqu'au bout, sans faillir. Merci pour tes nombreuses relectures et ton aide précieuse pour la traduction de certains textes anglais. Merci tout simplement d'être là.

A mes amis

Je remercie tous ceux qui m'ont encouragé et soutenu pendant ces trois années. Je pense tout particulièrement à Nicolas Boillot, fidèle ami, qui n'a jamais manqué de m'écouter avec attention et avec qui il est si agréable de discuter d'art, de science et des questions sur *la vie, l'univers et le reste*. Merci Nicolas de m'avoir donné goût au bavardage. Je pense également à Yvan Taurines, précieux ami, qui a toujours su, parfois avec maladresse, me donner son

¹ Bricolage au sens lévi-straussien du terme [*La pensée sauvage*, 1962], comme art de l'improvisation, de la spontanéité et de l'engagement.

point de vue lorsque parfois je manquais de raison. Merci Yvan pour ta franchise. Enfin, je remercie Gabriel, Manuel, Thomas, Isabelle, Julien et tous ceux qui m'ont apporté joie et distraction, lorsque la recherche prenait le pas sur ma vie privée, quand aussi, je n'en voyais plus le bout.

A l'Association Nationale de la Recherche et de la Technologie

Cette thèse a bénéficié du dispositif CIFRE (Convention Industrielle de Formation par la Recherche). Je suis très reconnaissant envers l'ANRT d'avoir accepté de subventionner l'entreprise Virtual-IT et de croire en la possibilité d'une collaboration juste et profitable entre laboratoires de recherche et entreprises. Merci d'avoir aidé mon projet de recherche à voir le jour.

Aux laboratoires de recherche : LARA et IRIT

Je remercie tous les acteurs du Laboratoire de Recherche en Audiovisuel (LARA) et de l'Institut de Recherche en Informatique de Toulouse (IRIT) de m'avoir accompagné tout au long de mon doctorat et d'avoir accepté la collaboration avec l'entreprise Virtual-IT. Je pense tout particulièrement à Guy Chapouillé, directeur de l'ESAV lorsque j'ai commencé mon activité de recherche. Merci Guy d'avoir créé cette école et de vous être battu pour défendre son identité. Nous portons nos rêves en concepts et nos concepts en images, en gardant en tête l'idée que vous avez toujours défendue, d'une praxis dont l'essence repose sur la réflexion, le sens critique, le « bricolage » et l'expérience.

AVANT-PROPOS

Cet avant-propos ne se substitue pas à l'introduction de ma thèse en ce sens où je souhaite en préambule rendre compte de mes motivations et des conditions de la réalisation de mon travail de recherche. Cet éclaircissement préliminaire donne un avant-goût de cet ouvrage et précise le cadre de travail qui m'a permis de venir à bout de cette grande et longue aventure.

Origine de la recherche

L'envie d'emprunter la voie de la recherche, à l'issue d'une formation initiale principalement technique, m'a été insufflée par le foisonnement créatif dans lequel j'ai évolué, issu de la rencontre entre l'art et la technologie, qui n'a eu de cesse de m'interroger sur les enjeux, les perspectives et les problématiques contemporaines liés au processus de création numérique, tant du côté de la conception (auteur) que de la réception (utilisateur). Impliqué à titre personnel et professionnel depuis une dizaine d'années dans la conception, le développement et le design d'interactions de projets numériques, je souhaitais témoigner, au travers de ma recherche, de la richesse extraordinaire issue du croisement entre la démarche scientifique et l'approche artistique. C'est à travers le DEA, étape préliminaire de ma recherche, que j'ai porté ma réflexion sur les installations numériques interactives.

Art et science sont complémentaires et partagent parfois des aspirations communes même si leur approche est souvent différente. « *L'artiste travaille du concept à l'analyse tandis que le scientifique travaille de l'analyse au concept.*² » Je dirais que le scientifique recherche la connaissance du monde en procédant pas à pas, en l'étudiant, afin de créer des modèles et des concepts alors que l'artiste propose d'abord un concept et tente de révéler une nouvelle vision du monde par l'expérience qu'il en propose. Art, science et technique entretiennent aujourd'hui un rapport privilégié et l'art numérique en est « exemplaire » à plus d'un titre car non seulement il est dans sa nature d'hybrider les techniques, qu'elles soient informatiques, électroniques ou optiques mais il favorise également la pluridisciplinarité par la collaboration entre scientifiques, ingénieurs et artistes qui y trouvent un terrain propice à la création et à l'expérimentation.

² KISSELEVA Olga. *Cyberart, un essai sur l'art du dialogue*. L'Harmattan. Paris, 1998, 368 p. (coll. L'Ouverture Philosophique)

L'implication des technologies informatiques et numériques pousse certains à refuser l'idée d'un art qui serait exclusivement dépendant de la machine et qui remettrait en cause le rôle de l'artiste et du spectateur vis-à-vis de l'œuvre. D'un autre côté, certains reconnaissent l'existence d'un « art technologique » sans pour autant nuancer leur discours en interrogeant la dimension artistique de la production contemporaine afin de distinguer ce qui relève de l'œuvre de ce qui relève du pur artifice technique, de la démonstration des possibles.

Ce sont ces préoccupations artistiques et scientifiques qui ont accompagné mon travail de recherche depuis ses débuts. Dans le cadre du DEA, j'ai tout naturellement orienté ma réflexion sur la question de la spécificité de l'art numérique ou « art des nouveaux médias ». En m'attachant à l'étude d'œuvres numériques projetant le spectateur dans une expérience interactive avec l'image, j'ai tenté de montrer comment l'œuvre numérique se déployait dans un espace de relation et de dialogue avec un « inter-acteur », ouvrant ou dilatant certaines dimensions esthétiques de l'image que je résume ici :

- La dimension du sensible, là où le corps du spectateur, par le biais de capteurs, est pleinement « actif » dans le déroulement de l'œuvre, dans son procès de production, en explorant l'espace, en créant un écart, une tension entre deux univers, celui de l'œuvre et celui de l'espace physique, là où « [...] *l'image actée relève à la fois d'une saisie de l'imaginaire et d'une saisie physique, interventionniste, par interfaces interposées : cette double détermination en constitue la singularité en regard de toutes les autres formes d'images.*³ »
- La dimension autopoïétique, lorsque la relation entre l'image interactive et le spectateur réunit les conditions nécessaires à son actualisation, à la révélation de sa morphogenèse et donc de la potentialité de l'œuvre. C'est en considérant la diamorphose de l'image numérique comme « moment esthétique », que sourd une éventuelle singularité, en considérant l'interactivité comme un moment où sont révélés les processus d'engendrement et de transformation de l'œuvre.

³ WEISSBERG Jean-Louis, *Présences à distance - déplacement virtuel et réseaux numériques - Pourquoi nous ne croyons plus la télévision ?*. L'Harmattan. Paris, 1999, 304 p. (coll. Communication et Civilisation)

- La dimension « interactive » comme lieu de rencontre, d'actualisation d'un espace-temps projectif, celui de l'entre-deux, issu de la rencontre entre le spectateur et l'œuvre. Une dimension entre l'espace réel et l'espace dématérialisé de la diégèse, ouverte à la participation comme un « *espace ménagé pour qu'un libre mouvement s'exerce*⁴. » Un univers simulé, régi par des lois algorithmiques, créé par la rencontre entre le corps du spectateur et son analogon numérique, l'image-créature, qui donne à nos sens l'illusion d'un monde sensible. L'espace de l'expérience sensible et singulière d'un « univers autonome », d'une fiction dont l'« écart », au sens aristotélicien du terme, est en permanence entretenu, à la fois par l'exploration du spectateur mais également par les règles de sa propre expérience.

Continuité de ma recherche

Dans la continuité de mes travaux de recherche et dans un souci de rapprochement avec le milieu professionnel et industriel, j'ai pris contact avec l'entreprise *Virtual-IT* qui a trouvé dans mes travaux de recherche un intérêt à monter une collaboration sous la forme d'un dispositif CIFRE (Conventions Industrielles de Formation par la Recherche). Cette thèse est donc le fruit d'une collaboration soutenue entre l'entreprise Virtual-IT, le Laboratoire de Recherche en Audiovisuel (LARA) et l'Institut de Recherche en Informatique de Toulouse (IRIT). La société Virtual-IT s'est avérée être un partenaire idéal, tant dans son engagement vis-à-vis de mon projet de recherche (moyens financiers et humains) que dans l'accueil irréprochable qui m'a été réservé. Dès mon arrivée dans l'entreprise, nous avons très vite trouvé un terrain d'entente sur les directions de la recherche que nous allions suivre. Depuis quelques années déjà, je m'intéressais particulièrement à la manière dont les artistes s'appropriaient la technologie pour proposer aux spectateurs de nouvelles modalités d'interaction avec l'œuvre. Impliqué dans de nombreux projets d'installations numériques, et proche de laboratoires d'informatique comme InfoMus de l'université de Gênes, j'avais la volonté de comprendre les enjeux des nouvelles technologies de captation du corps et des gestes dans la pratique artistique. Quelques temps avant le début de mon doctorat, je tombai sur une vidéo présentée par Jeff Han, chercheur à l'université de New York, qui faisait la démonstration d'un écran translucide et multi-tactile. L'effet a été quasiment immédiat. En faisant le constat que cette technologie sortait enfin du domaine exclusif de la

⁴ BOISSIER Jean-Louis. *La relation comme forme - L'interactivité en art*. Mamco. Genève, 2004, 311 p.

recherche et des arts, je pris contact avec Jean-François Pal, directeur de Virtual-IT, pour lui proposer de travailler dans cette direction. Les interfaces multi-tactiles ajoutaient une nouvelle dimension à l'image numérique, celle du toucher et de l'exploration par le geste, dont la nature exacte me restait encore à découvrir...

Cadre professionnel

Afin de répondre à la stratégie générale de l'entreprise et à la volonté d'orienter son activité vers l'innovation, j'ai eu la chance d'avoir rapidement la responsabilité de la création d'un pôle de recherche et développement dont les principaux axes se sont concentrés sur l'élaboration de dispositifs tactiles interactifs, la définition d'un modèle d'interaction et la création d'un environnement de développement multi-tactile gestuel. Depuis, de nombreux acteurs de l'innovation nous ont rejoint ou nous accompagnent dans nos projets. Nous avons en effet le soutien de Midi-Pyrénées Innovation (MPI), nous avons lancé un partenariat depuis janvier 2009 avec l'école d'ingénieurs Epitech et d'autres entreprises du tissu économique régional et national nous accompagnent dans nos projets. Les efforts que nous avons fournis depuis trois ans nous ont permis d'obtenir des résultats et des développements concrets dans le domaine des technologies tactiles multipoint, des avancées qui nous ont rapprochés de clients comme les groupes DCNS, THALÈS, INEO / SUEZ ou encore TOTAL. Afin d'évaluer nos technologies et nos hypothèses de travail, nous avons participé activement à des salons comme Milipol Paris, Laval Virtual, Euronaval et d'autres, davantage « grand public », comme le TGS (*Toulouse Game Show*). Les retours que nous avons eu et les observations faites sur le terrain, notamment par le biais d'enregistrements vidéo et logiciels, nous ont permis de progresser rapidement dans la validation de nos hypothèses de recherche et la définition d'un modèle d'interaction.

La thèse, juste un commencement

Aux premiers temps de la thèse, j'avoue qu'il m'était assez difficile d'envisager comment concilier mes propres velléités de recherche, les impératifs et les contraintes industrielles de l'entreprise, et la nature de la thèse qui suppose de proposer un regard original et pertinent sur le sujet de recherche. Je crois que l'erreur que je commettais alors était d'envisager la thèse comme un aboutissement, voire une fin en soi, et c'est donc avec humilité et la plus

grande modestie que j'ai appris à mener mon travail de recherche. Le sentiment d'amertume qui m'avait envahi laissa progressivement place à la joie lorsque je compris que la thèse n'était qu'un commencement.

Cependant, au fil des multiples recherches documentaires que j'ai menées, une nouvelle frustration naissait progressivement. Les idées que je pensais originales et pertinentes au départ et certaines directions de recherche que j'envisageais se retrouvaient chacune, déjà formulées et largement couvertes. Ce n'est qu'après de longues journées de recherche et de longues nuits de développement que les premiers indices nécessaires aux hypothèses que je défends dans ma thèse ont commencé à apparaître. Ceci a considérablement facilité ma recherche documentaire en m'éclairant sur les pistes théoriques que je devais suivre et les points restés sous silence ou pas suffisamment analysés, qui allaient me permettre d'établir l'angle d'approche de mon analyse. Suite à ces trois années, je m'aperçois que ce doctorat ne représente qu'un travail préliminaire, des portes ouvertes vers de multiples questions, des chapitres dont le sujet seul mériterait qu'on lui consacre une thèse. Il me tarde déjà de poursuivre le travail engagé...

Cette préface, ou « *salsa del libro* » (sauce du livre) comme la nomme les italiens, m'a permis de présenter l'origine de mes travaux de recherche, le cadre de la réalisation de ma thèse et les facteurs techniques et humains qui ont participé à son bon déroulement. J'espère que ce préambule vous a suffisamment mis en appétit pour en continuer la lecture.

SOMMAIRE

Une table des matières détaillée est disponible à la fin de l'ouvrage

REMERCIEMENTS	3
AVANT-PROPOS	7
SOMMAIRE	12
INTRODUCTION.....	14
CHAPITRE I CHAMP DE RECHERCHE.....	26
1. INTERACTION HOMME-MACHINE	31
2. INSTRUMENT D'INTERACTION.....	40
CHAPITRE II INTERFACES TACTILES MULTI-POINTS.....	64
INTRODUCTION	65
1. UNE LONGUE PERIODE D'INCUBATION	68
2. HISTOIRE DES TECHNOLOGIES MULTI-TACTILES	70
3. PRE-APPROPRIATION TECHNOLOGIQUE	89
CHAPITRE III ESPACE DE CONCEPTION.....	101
1. DOMAINES D'APPLICATION	103
2. PERIPHERIQUES D'ENTREE MULTI-TACTILES.....	104
3. PERIPHERIQUES DE SORTIE.....	115
4. LE CANAL GESTUEL	117
CHAPITRE IV MODÈLE D'INTERACTION	135
1. INTERFACE UTILISATEUR NATURELLE.....	138
2. MODELISATIONS TEMPORELLES DU GESTE.....	146

3.	MODELISATION SPATIALE DU GESTE.....	149
4.	MODALITES DES INTERACTIONS.....	157
5.	TAXINOMIE DU GESTE	165
6.	DESIGN D'INTERACTION	173
 CHAPITRE V PROTOTYPES		186
1.	PROTOTYPAGE	187
2.	PREREQUIS LOGICIELS	198
3.	CONSIDERATIONS TECHNIQUES	203
 CHAPITRE VI ENVIRONNEMENT DE DÉVELOPPEMENT.....		209
1.	EXTENSION DU MODELE D'INTERACTION	211
2.	DEVELOPPEMENT DE L'INTERGICIEL MULTI-TACTILE.....	231
3.	IMPLEMENTATION DES API	259
 CONCLUSION.....		268
 ANNEXES		284
1.	ANNEXES GENERALES	285
2.	DESIGN ET CONCEPTS ART.....	289
3.	SALONS ET DEMONSTRATIONS	295
 BIBLIOGRAPHIE		301
TABLE DES ILLUSTRATIONS		310
LISTE DES TABLEAUX		313
TABLE DES MATIÈRES		314

INTRODUCTION

« Man becomes as it were the sex organ of the machine world, as the bee of the plant world, enabling it to fecundate and to evolve ever new forms. »

« L'homme devient l'organe de reproduction du monde des machines, comme l'abeille celui des plantes, autorisant sa fécondation et son évolution vers de nouvelles formes. »

Marshall McLuhan

Le système technicien

Depuis la révolution industrielle, les « systèmes techniques », définis par l'historien Bertrand Gille comme des ensembles cohérents où s'articulent les différentes technologies d'une époque, ont influencé de manière considérable les systèmes économiques et sociaux. La technique s'est infiltrée dans les sphères industrielles sous des formes multiples, comme moyen d'accélération des processus de production, comme méthode de contrôle des cycles de vie des produits, comme vecteur de communication et comme l'objet même de la production. Le développement des machines, et plus particulièrement celui des machines informatiques, a marqué un pas supplémentaire dans la pénétration des objets techniques dans notre environnement. En affirmant que « *l'homme n'est rien d'autre que la vie technique*⁵ », Bernard Stiegler radicalise la pensée de Gille et suggère que le détachement progressif des cycles technologiques et la pénétration excessive des objets techniques dans notre vie nous place dans une situation où nos facultés de compréhension et d'appropriation sont altérées. Il place son analyse en marge de la tradition grecque en renversant la proposition selon laquelle la technique serait au service de l'homme en la décrivant comme « *l'horizon de toute possibilité à venir et de toute possibilité d'avenir* ». Stiegler définit ainsi la technique « contemporaine » non pas « *en termes de fins et de moyens* » [Aristote], mais l'aborde en soulignant les liens étroits entre *tekhnè* et *épistémè* (technoscience), qui sont selon lui à l'origine de l'accélération dans la succession des « *ruptures technologiques*⁶ ».

La technique semblerait n'obéir qu'à sa propre logique et disposer d'une certaine autonomie, échappant progressivement à l'emprise des hommes. Dès 1954, Jacques Ellul, dans *La technique ou l'enjeu du siècle*, dresse un portrait peu réjouissant du système technicien en partageant ses craintes que l'homme ne soit « attaqué », « atteint profondément dans ses sources vitales ». La technique modifierait notre prise au réel, créant une société abstraite et homogénéisée, une « société virtuelle ». Ellul cite Jean Beaudrillard [Société de consommation, 1970] en évoquant la difficulté croissante de considérer la rationalité des objets de la consommation « *car la technique efface le principe même de réalité (sociale)* ». Gilbert Simondon déclare quant à lui que l'objet industriel, contrairement à l'objet artisanal, reçoit moins ses caractères de l'extérieur que des contraintes de sa genèse, de son processus de concrétisation et du phénomène généralisé de standardisation.

⁵ STIEGLER Bernard. *La technique et le temps - La faute d'Epiméthée* (Tome 1). Galilée. Paris, 1994, 284 p.

⁶ GILLE Bertrand (Ouvrage collectif sous la direction de). *Histoire des techniques : Technique et civilisations, technique et sciences*. Paris, Gallimard, 1978, 1652 p. (Coll. Encyclopédie de La Pléiade)

L'espace accordé à l'intentionnalité humaine est diminué par les logiques économiques et par la production de masse, privant ainsi les utilisateurs de toute participation critique à l'évolution des produits. Ainsi, Gilbert Simondon décrit le système technicien en déclarant que « *les besoins se moulent sur l'objet technique industriel, qui acquiert ainsi le pouvoir de modeler une civilisation.*⁷ » Néanmoins, il nous est permis de penser qu'il existe encore des lieux où l'homme peut garder le contrôle par l'« anticipation » des systèmes techniques, c'est notamment ceux de la praxis artistique, de l'invention et de l'expérimentation. En explorant ces espaces de critique et de liberté, il me semble qu'il demeure toujours possible d'avoir conscience de l'inscription de notre « temporalité » dans celle du système technicien.

Design centré utilisateur

« *De nos jours, c'est l'individu qui se conforme aux besoins de la technologie. Il est temps que la technologie se conforme aux besoins de l'individu.*⁸ » Par cette déclaration de 1988, Donald Norman cristallise des concepts déjà au cœur des préoccupations d'une poignée de chercheurs en interaction homme-machine et dont les premiers travaux remontent aux années soixante-dix. C'est notamment au sein du centre de recherche PARC Xerox qu'apparaissent les concepts d'interface utilisateur graphique et de manipulation directe. Ces principes convoquent l'idée essentielle que les technologies informatiques, au cours du procès de leur développement, devraient impliquer davantage les besoins, les attentes et les retours de l'utilisateur. Cette période correspond à la formalisation du principe de design centré utilisateur qui vise à intégrer, à chaque phase du développement d'un produit, la satisfaction, les compétences et les limites physiques et cognitives de l'utilisateur. Dès lors, la recherche en interaction homme-machine n'a eu de cesse d'emprunter et de croiser les champs de la psychologie, des sciences cognitives, de l'anthropologie et des arts, pour améliorer l'efficacité et l'usabilité des systèmes informatiques.

Dans le prolongement de cette réflexion, Donald Norman [*The design of everyday things*, 1998] et John Maeda, [*The Laws of Simplicity*, 2006] défendent l'idée que la complexité d'un système devrait rester « invisible » à l'utilisateur afin de ne pas nuire à la simplicité de son

⁷ SIMONDON Gilbert. *Du mode d'existence des objets techniques*. Aubier, Paris, 2001, 333 p.

⁸ « Today, it is the individual who must conform to the needs of technology. It is time to make technology conform to the needs of people. » In NORMAN Donald Arthur. *The Invisible Computer : why good products can fail, the personal computer is so complex, and information appliances are the solution*. MIT Press, Cambridge, septembre 1998, 316 p. Traduction libre : COQUELIN Marine et MORIN Rudy.

utilisation. Malgré cette conviction substantielle et largement partagée par les designers, il est aisé de constater que paradoxalement, les industriels, piégés par des cycles de développement produit de plus en plus rapides et conduits par des logiques de concurrence de marchés, semblent accorder davantage d'intérêt à l'amélioration technique de leur produit qu'au respect de ce principe de « simplicité ». Il suffit de s'attarder sur les objets technologiques qui nous entourent pour discerner que l'objectif affiché n'est pas tant de nous simplifier la tâche mais plutôt d'en réaliser le plus possible. Dans cette surenchère technique, l'homme ne dispose que d'un temps très court pour s'adapter à ces technologies, en apprécier leur juste valeur et participer à leur processus de conception et d'évolution, comme semblerait le suggérer le mythe idéaliste du design centré utilisateur revendiqué par certains industriels. Ce que nous savons généralement de la nature des objets techniques que nous utilisons pourtant quotidiennement, concerne davantage leurs caractéristiques, dont l'évolution est rythmée par une communication excessive sur leurs nouvelles fonctionnalités, que les formes et les enjeux de leur utilisation. Privés des repères pour évaluer leur qualité et leur simplicité, les utilisateurs consomment avec démesure les technologies informatiques sans réellement s'intéresser à leur fonctionnement et moins encore aux perspectives de leur évolution.

Origine de la recherche

Dans un contexte où les industriels décrivent les technologies en vantant leurs performances ou en énumérant la liste de leurs applications possibles, l'utilisateur a des raisons de croire que les vertus d'un produit dépendent de mesures quantitatives, le privant ainsi des moyens d'avoir si ce n'est un regard critique, une appréciation plus juste de sa valeur. Cette frustration est je pense à l'origine de la motivation qui a porté la rédaction de cette thèse et les développements qui en sont issus. En effet, le début de ma recherche en 2007 coïncide avec la présentation de la table interactive *Microsoft Surface* et la commercialisation de la première version de l'*iPhone* d'*Apple*, dont on nous vantait, à grand renfort de marketing, que ces produits disposaient d'un écran multi-tactile, en ne sachant pas comment nous en montrer l'intérêt, si ce n'est par la démonstration, devenue ridicule, de la gestion d'un album photo. Conscient des enjeux soulevés par ces technologies et vouant déjà un intérêt tout particulier aux interfaces utilisateur tangibles ou multi-tactiles, qu'il s'agisse de

contributions scientifiques (Jeff han⁹ 2006, Mark Weiser¹⁰) de formes artistiques singulières (*ReacTable*¹¹, 2005) ou d'autres projets expérimentaux (*Khronos Projector*¹², 2005), je ne pouvais me résoudre à voir quarante années de recherche réduites à un argument marketing, sans m'engager personnellement dans leur étude et participer à l'évolution du design d'interaction.

Je ferais volontiers l'aveu que cette recherche n'est point la conséquence d'une obsession naïve pour les artifices techniques, mais qu'elle relève davantage du besoin que j'éprouve de rationaliser les objets technologiques qui nous entourent et de comprendre en quoi ils modifient notre sensibilité. En rejoignant les visions dépeintes par les travaux scientifiques, les œuvres littéraires, audiovisuelles et artistiques qui appartiennent à l'imaginaire collectif, les interfaces utilisateur multi-tactiles ont fixé mon attention et orienté ma réflexion. En se parant de la dimension du sensible par le truchement de capteurs et se débarrassant progressivement de son « corps » physique par la suppression des périphériques utilisateur, ces dispositifs tendent à s'effacer devant l'image interactive, plus présente que jamais. L'interactivité devient progressivement une disposition naturelle de leur « existence ». La curiosité et l'intuition m'ont alors amené, aux prémices de ma recherche, à me poser la question suivante :

Quelle est la nature de la relation homme-machine impliquée dans les interfaces utilisateur multi-tactiles ? Et plus précisément, quelles en sont les modalités d'interaction ?

En s'inspirant des modalités de la perception, de l'action et de la communication de l'homme, les chercheurs en IHM manifestent l'ambition de créer les conditions d'une relation homme-machine plus « naturelle ». Les gants de données ou les combinaisons bardées de capteurs, les caméras tridimensionnelles, les capteurs de mouvement et de pression et les accéléromètres sont autant d'instruments qui donnent aux ordinateurs les moyens d'être « sensibles » à leur environnement. Les interfaces de ces dispositifs,

⁹ HAN Jefferson Y. *Low-cost multi-touch sensing through frustrated total internal reflection*. In *Proceedings of the 18th annual ACM symposium on User interface software and technology* (UIST '05). ACM, New York, USA, 2005, pp. 115-118

¹⁰ WEISER Mark. *The computer for the 21st century*. In *Human-computer interaction*, Ronald M. Baecker, Jonathan Grudin, William A. S. Buxton, and Saul Greenberg. San Francisco, USA, Weiser, pp. 933-940

¹¹ REACTABLE. <http://www.reactable.com>

¹² CASSINELLI Alvaro. *The khronos projector. A video time-warping machine with a tangible deformable screen*. Disponible ici : <http://www.k2.t.u-tokyo.ac.jp/members/alvaro/Khronos/>

autorisant des interactions multimodales, agissent comme des lieux d'exploration et d'actualisation, où le corps sensible de l'utilisateur se heurte aux images, espace de la représentation, et aux capteurs, organes sensoriels d'une machine anthropo-mimétique. L'esprit « dé-joue » les lois d'un monde algorithmiques, engendrant les conditions d'un contact, d'une relation et d'un dialogue. Selon Edmond Couchot, « *cette ouverture sur le monde environnant, associée à des possibilités d'actions physiques sur ce même monde, modifie en profondeur la relation traditionnelle de ce que le philosophe Georges Simondon appelle le « couplage » entre l'homme et la machine.*¹³ »

Le déplacement conceptuel de la notion d'instrument d'interaction causé par l'abstraction du périphérique d'entrée et sa « ré-in-carnation » dans le geste inscrivent les interfaces tactiles dans les concepts de l'informatique ubiquitaire, en ce sens où la disparition physique du périphérique d'entrée contribue à l'« intégration invisible » des technologies informatiques dans notre environnement. En s'affranchissant de la contrainte mécanique des périphériques de saisie, par le glissement de la partie physique de l'instrument vers la main, les interfaces tactiles multi-points créent des conditions propices à l'émergence de nouvelles interactions.

Etat de la recherche

Depuis les textes précurseurs de Lloyd Nakatani et John Rohrlich¹⁴, ingénieurs aux laboratoires *Bell* dans les années 80, ayant introduit la notion de « *machines flexibles* » (ou *soft-machines*) pour qualifier la souplesse et la modularité des contrôleurs graphiques (*soft-controls*) comparées à celles des périphériques physiques, de nombreuses études se sont intéressées aux interfaces tactiles afin de mesurer leurs avantages par rapport aux interfaces WIMP traditionnelles. Les contributions antérieures de Card et al.¹⁵, reprises par Sears et Shneiderman¹⁶ en 1991, ont démontré le gain de performance et la réduction

¹³ COUCHOT Edmond. *Des changements dans la hiérarchie du sensible*. In *Les cinq sens de la création*, BORILLO Mario et SAUVAGEOT Anne. Champ Vallon, novembre 1998, 220 p. (Coll. Milieux)

¹⁴ NAKATANI Lloyd H. et ROHRlich John A. *Soft machines : A philosophy of user-computer interface design*. In *Proceedings of the SIGCHI conference on Human Factors in Computing System (CHI '83)*. ACM, New York, USA, 1983, p. 19-23

¹⁵ CARD Stuart K., ENGLISH W. K. et BURR B. J. *Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys, for text selection on a CRT*. In *Human-computer interaction*, BAECKER R. M. Morgan Kaufmann Publishers Inc., San Francisco, USA, 1987, pp. 386-392

¹⁶ SEARS Andrew et SHNEIDERMAN Ben. *High precision touchscreens : design strategies and comparisons with a mouse*. In *International journal of man-machine studies*, 1991, pp. 593-613

d'erreurs de pointage conférés par les interactions directes dans le cas d'objets graphiques dont la taille est suffisante. Des travaux¹⁷ plus récents ont apporté des solutions pour remédier aux difficultés de sélection d'objets de taille inférieure à celle du doigt. Les études dans le domaine des interactions bi-manuelles, notamment les approches théoriques de Guiard et Ferrand¹⁸ et les études empiriques de Buxtons et Myers¹⁹ ou Fitzmaurice et al.^{20 21}, ont clairement démontré les avantages de l'interaction bi-manuelle permise exclusivement par des interfaces multiplexées dans l'espace.

La majorité des études menées jusqu'à présent fait état des différences de performance comparées entre les interfaces tactiles et les interfaces traditionnelles, mais très peu se sont attachées à intégrer dans leur approche les spécificités conceptuelles et techniques de l'interaction tactile multi-points. Ces travaux ne sont que rarement mis en perspective avec la technologie d'acquisition employée, la configuration et la taille du dispositif. Enfin, les travaux réalisés dans le domaine des interfaces multimodales, notamment les directions de recherche combinant le langage naturel et le geste, ont intégré dans leur approche des éléments d'analyse issus des domaines de la psychologie et de la linguistique [Piaget, McNeill, Nespoulos et Lecours], permettant d'obtenir les taxinomies générales du geste en situation d'interaction homme-machine [Quek et al.] [Karam et Schraefel]. Néanmoins, ces classifications me semblent trop générales pour caractériser de manière précise les interactions multi-tactiles.

¹⁷ BENKO Hrvoje, WILSON Andrew D. et BAUDISCH Patrick. *Precise selection techniques for multi-touch screens*. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (CHI '06), GRINTER Rebecca, RODDEN Thomas, AOKI Paul, CUTRELL Ed, JEFFRIES Robin et OLSON Gary. ACM, New York, USA, 2006, pp. 1263-1272

¹⁸ GUIARD Y. et FERRAND T. *Asymmetry in bimanual skills*. In *Manual asymmetries in motor performance*, ELLIOTT D. et ROY E.A. CRC Press, 1995, pp. 175-195

¹⁹ BUXTON William et MYERS Brad A. Study in two-handed input. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (CHI '86), Marilyn Mantei and Peter Orbeton. ACM, New York, NY, USA, 321-326. 1986

²⁰ KURTENBACH Gordon, FITZMAURICE George, BAUDEL Thomas et BUXTON Bill. *The design of a GUI paradigm based on tablets, two-hands, and transparency*. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (CHI '97). ACM, New York, USA, 1997, pp. 35-42

²¹ OWEN Russell, KURTENBACH Gordon, FITZMAURICE George, BAUDEL Thomas et BUXTON Bill. *When it gets more difficult, use both hands: exploring bimanual curve manipulation*. In *Proceedings of Graphics Interface* (GI '05). Canadian Human-Computer Communications Society, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 2005, pp. 17-24

Méthodologie de la recherche

Pour apporter des éléments de réponse à la question initiale et rendre compte de la nature des modalités d'interaction homme-machine impliquées dans les interfaces utilisateur multi-tactile, j'articulerai cette étude en faisant l'hypothèse que les interfaces tactiles sont des formes singulières d'interface gestuelle en ce sens où la manipulation d'objets graphiques par les sens tactilo-kinesthésiques suppose la participation limitée, puisqu'en deux dimensions, des notions innées de l'exploration et de la proprioception. De plus, la gestuelle abstraite de certaines commandes systèmes, par la réalisation de mouvements déictiques, symboliques ou pictographiques, tend à imiter les modalités de la communication entre les hommes, sans nécessairement prendre en compte le caractère culturel et subjectif des gestes ni les difficultés liées à leur apprentissage. En partant du postulat que les interfaces tactiles multi-points convoquent un type de relation spécifique et profondément lié à la nature du geste et de la technologie, je défends l'idée que le design d'interface multi-tactile ne saurait être approché qu'en suivant un modèle d'interaction intégrant l'ensemble des composantes physiques, cognitives, sensorielles et motrices du geste dans le couplage homme-machine.

Dans la perspective d'élaborer un modèle d'interaction multi-tactile gestuel, j'articulerai ma recherche autour des problématiques suivantes :

- En quoi les caractères du « geste à nu » et de la main déterminent-ils les modalités de couplage avec le système ?
- En quoi la morphologie des dispositifs multi-tactiles influence-t-elle le design d'interaction ?
- Dans quelle mesure le geste, impliqué dans les interactions multi-tactiles, peut-il être qualifié de « naturel » ?
- Quelles sont les caractères et les limites de l'imitation du réel, du réalisme des interactions de manipulation ?
- Dans quelle mesure les critères de visibilité, de mapping naturel, d'affordance et de feedback améliorent-ils la « simplicité » de ces interfaces ?

Ma méthodologie de recherche reposera sur une approche hypothético-déductive, au croisement des sciences informatiques, esthétiques, anthropologiques et psychologiques afin de profiter de la richesse des contributions pluridisciplinaires qui caractérise la recherche en IHM. En prolongeant les travaux existants, je serai en mesure de définir un modèle théorique prédictif, dont les concepts seront appliqués et mis à l'épreuve par des applications concrètes. Ainsi, mon étude profitera des phases de développement et des retours d'expériences utilisateur pour valider et enrichir le modèle théorique défini préalablement.

Annonce du plan

Au chapitre I, nous verrons quels sont les ressorts et les fondements théoriques qui ont accompagné les concepts précurseurs de la recherche en IHM à travers les contributions scientifiques déterminantes de Douglas Engelbart, Ivan Sutherland, Alan Kay et Ben Shneidermann. J'apporterai des éléments de compréhension indispensables à l'approche du paradigme WIMP, défini dès 1980, et qui demeure aujourd'hui encore le modèle dominant. Pour apporter un éclairage sur la nécessité et les enjeux de la recherche engagée en IHM, je reviendrai sur les principes d'interaction définis par la « manipulation directe », thème transversal aux directions théoriques et pratiques abordées dans cet ouvrage. Je placerai donc l'analyse de mon objet d'étude, les interfaces multi-tactiles, dans une continuité de la recherche, où les notions d'interface graphique, de périphérique d'entrée et d'interaction instrumentale, serviront de substrat à la compréhension du changement de paradigme amené par les interfaces tactiles multi-points.

Nous nous pencherons également sur deux formes d'interface post-WIMP qui nous permettront d'identifier l'éventuelle spécificité de notre objet d'étude. Dans un premier temps, mettre en exergue l'approche proposée par les interfaces haptiques nous donnera un cadre de compréhension pour apprécier la richesse des perceptions tactiles et kinesthésiques impliquées dans le geste épistémique. En apportant un éclairage sur les concepts d'interface tangible et les moyens d'action permis par les interacteurs mixtes, sorte d'artefacts physiques aux propriétés réelles et virtuelles, nous aborderons par l'exemple des principes qui participeront à la définition du geste ergotique. Enfin, l'étude de ces dispositifs nous permettra de relever en quoi ces différentes formes s'inscrivent en des points distincts du continuum de virtualité, principalement par leur appropriation singulière de la notion

d'instrument, du geste prothétique au « geste à nu », afin d'ouvrir notre étude sur l'analyse approfondie des caractères du geste manuel impliqués dans les interfaces multi-tactiles.

Le chapitre II me permettra de présenter mon objet d'étude des années soixante-dix à nos jours, à travers son évolution conceptuelle et technique, en mettant en lumière la longue période d'innovation où se sont succédés l'invention, l'affinage technique et l'adoption de cette technologie par les industriels et les utilisateurs. D'une part, nous chercherons à identifier les concepts rattachés à cette technologie, les raisons de leur appropriation précoce par les utilisateurs et le stade actuel de leur évolution. D'autre part, nous comprendrons de quelle manière l'art numérique, le jeu vidéo, la science-fiction et les communautés de chercheurs amateurs ont participé aux réflexions et à l'émergence des modalités d'interaction tactiles et gestuelles.

Le chapitre III s'intéressera à définir un espace de conception qui permettra de clarifier l'analyse des technologies multi-tactiles et servira d'approche théorique et technique à la définition du modèle d'interaction. Nous prolongerons la classification sur les IHM gestuelles proposée par Maria Karam et M.C. Schraefel²² en réalisant une taxinomie spécifique à notre objet d'étude. Dans un premier temps, nous ferons une présentation rapide de ses domaines d'application et des aspects généraux de sa morphologie. Cette partie présentera également les interfaces multi-tactiles sous un angle technique, voire physiologique, en décrivant les différentes technologies d'acquisition, leurs qualités et leurs défauts. Les éléments techniques que nous aurons relevés serviront, dans le cadre de la réalisation des prototypes multi-tactiles, à définir les orientations et l'échéancier du programme de recherche et développement. Enfin, nous aborderons le canal gestuel, en proposant une définition préliminaire avant de préciser progressivement les caractères du geste impliqués dans les interactions multi-tactiles. Nous suivrons la typologie fonctionnelle proposée par Claude Cadoz, en distinguant le geste épistémique, ergotique et sémiotique en vue d'obtenir une taxinomie générale du geste.

Le chapitre IV sera consacré à la définition d'un modèle d'interaction qui constituera un cadre de compréhension théorique à l'évaluation, la comparaison et la création de design d'interaction multi-tactile gestuel. Nous profiterons de cette partie pour préciser les limites du terme « naturel » dans « interface utilisateur naturelle », en revenant sur la confusion qui

²² KARAM Maria et Schraefel M.C. *A taxonomy of gestures in human computer interaction*. In *Transactions on Computer-Human Interactions*, TBD, 2005

entoure sa définition et prévenir toute interprétation abusive. En nous appuyant sur le paradigme OCGM (*Objects, Containers, Gesture, Manipulation*) proposé par Ron George et Joshua Blake²³, nous présenterons l'approche largement répandue en design d'interaction multi-tactile, qui consiste à utiliser la métaphore du réel pour représenter l'objet de la manipulation et son actualisation. Je ferai part également dans ce chapitre de l'intuition qui m'a accompagnée depuis les débuts du projet de thèse, remettant en cause le principe de simulation du réel, en posant les limites du réalisme dans l'efficacité de tels systèmes. Je reviendrai notamment sur ce point, en inscrivant ma recherche dans les théories de la perception défendues par Donald Norman²⁴ qui présentent un point de vue complémentaire à l'écologie de la perception proposée par James Gibson²⁵. J'effectuerai la distinction conceptuelle entre l'affordance réelle et l'affordance perçue des objets graphiques conférée par le réalisme de leur représentation et de leur comportement. Le reste du chapitre décrit la typologie formelle du geste, la modélisation spatiale et temporelle de la phrase gestuelle et les modalités d'interaction conditionnées par la morphologie des dispositifs multi-tactiles. Enfin, nous aborderons le design d'interaction sous l'angle de la « simplicité », en nous appropriant les principes de visibilité, de mapping et de feedback qui accompagneront le développement des projets réalisés.

Les chapitres V et VI rendront compte de l'inscription de cette étude dans la démarche d'une recherche appliquée, sous la forme d'un dispositif CIFRE, convention industrielle de formation par la recherche, dont le financement a autorisé mon projet de thèse à voir le jour. La collaboration efficace entre les laboratoires de recherche du LARA, de l'IRIT et l'entreprise Virtual-IT, m'a permis de réaliser les dispositifs et les études nécessaires à l'expérimentation et à la mise en pratique des directions de recherche avancées dans la thèse. Dans un premier temps, je présenterai les prototypes que nous avons conçus en exposant les raisons qui nous ont conduits à privilégier certaines configurations puis je détaillerai les considérations techniques issues de leur expérimentation. Ensuite, nous nous attarderons sur la méthodologie de recueil et de traitement des données des phases d'observation qui nous ont permis d'« enrichir » le modèle d'interaction théorique. Enfin, je détaillerai la conception et le développement de l'environnement multi-tactile (*framework*),

²³ GEORGE Ron et BLAKE Joshua. *Objects, Containers, Gestures, and Manipulations: universal foundational metaphors of natural user interfaces*. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (CHI '10), Atlanta, Georgia, USA, 2010

²⁴ NORMAN Donald Arthur. *Affordances, conventions, and design*. In *Interactions*, Volume 6 Issue 3, 1999, pp. 38-42

²⁵ GIBSON James J. *L'approche écologique de la perception visuelle*. MF, 2010, 332p. (Coll. Dehors)

constituant mon projet de thèse, en insistant sur la logique de conception et l'architecture logicielle dont les principes sont directement issus des directions ouvertes par le modèle d'interaction préalablement défini.

La conclusion me permettra de revenir sur les résultats de cette étude, en prenant davantage de distance critique puisque mon projet de thèse sera abouti. Je profiterai donc de ce moment pour revenir sur les résultats de la recherche et présenter une perspective des travaux à venir. D'une part, j'aborderai des points techniques en précisant les limites de l'environnement de développement multi-tactile existant et en abordant les développements que nous envisageons. D'autre part, ce chapitre s'inscrira dans la thèse comme un point d'orgue théorique, en ce sens où je coucherai sur le papier certaines intuitions qui, à mesure que ma thèse avançait, se sont cristallisées en convictions. En prolongeant les pistes théoriques soulevées au chapitre IV sur les limites de la représentation réaliste des interactions tactiles, je tenterai de montrer qu'en plaçant le design d'interface et d'interaction uniquement sous le joug d'une simulation du réel, la relation homme-machine ne risque de prendre davantage les allures d'un simulacre.

CHAPITRE I CHAMP DE RECHERCHE

« La science informatique n'est pas plus la science des ordinateurs que l'astronomie n'est celle des télescopes »

«Computer science is no more about computers than astronomy is about telescopes.»

Edsger Dijkstra

Le champ de recherche relatif aux interactions homme-machine convoque aujourd'hui la contribution pluridisciplinaire des sciences informatiques, de la psychologie, de l'anthropologie, de l'éducation, du design, des mathématiques, de la physique et de l'esthétique. L'apparition et l'évolution de l'interface graphique utilisateur sont directement liées à l'intégration des facteurs humains dans la relation entre l'utilisateur et la machine. Douglas Engelbart prend conscience très tôt de l'importance à considérer le contexte social et culturel de l'homme et donne un nouveau cadre de compréhension aux notions d'utilisateur et d'ordinateur personnel. La contribution d'Alan Kay est également considérable dans l'avènement de l'interface graphique utilisateur et les codes perceptuels qui lui sont rattachés inspirés des travaux sur la psychologie du développement et l'acquisition du savoir menés par les psychologues Jean Piaget et Jerome Bruner.

A travers ce chapitre, nous verrons comment l'étude des processus de la perception et de la cognition s'est révélée incroyablement féconde pour le design de systèmes interactifs, pour analyser et limiter les phénomènes d'erreurs et donner plus de satisfaction à l'utilisateur. Comprendre le passage du modèle d'interaction WIMP aux modèles POST-WIMP implique une meilleure compréhension des processus engagés dans la relation homme-machine, en tenant compte des compétences humaines comme la perception, la cognition et l'action. Nous chercherons à comprendre comment certains modèles d'interaction se sont imposés afin d'obtenir un cadre de compréhension à l'émergence de nouvelles interfaces utilisateur.

Les interfaces post-WIMP représentent une voie de sortie du paradigme WIMP en ce sens où elles proposent de nouveaux moyens d'interaction et de manipulation des objets graphiques. Les recherches dans le domaine des IHM explorent de nouvelles techniques d'interaction afin de créer des interfaces plus « naturelles » en créant de nouveaux périphériques utilisateur, en ajoutant de nouveaux canaux de communication homme-machine par le biais de capteurs électroniques et optiques. Dans le but d'« humaniser » la relation homme-machine, les chercheurs se sont penchés sur la création d'interfaces et de modes dialogiques qui intègrent davantage les caractéristiques naturelles de la perception et de la communication humaine.

Une des voies de la recherche issue de ces préoccupations sont les interfaces multimodales. Ces systèmes proposent plusieurs modalités de captation de l'activité de l'utilisateur. Ils coordonnent plusieurs types d'entrées, comme la parole, les gestes, les expressions faciales

ou les mouvements pour analyser et interpréter les actions de l'utilisateur. Ce type d'interface a pour fonction de reconnaître les formes naturelles du langage et du comportement humain afin d'étendre le potentiel de communication des hommes avec la machine. Dans le même temps, l'apparition des gants de données (*datagloves*), des combinaisons bardées de capteurs et l'avènement de la réalité virtuelle, ont placé très tôt le geste comme un canal de communication privilégié. Comme le souligne Claude Cadoz, le geste est peut-être « *le plus singulier et le plus riche des canaux de communication.* »²⁶ En outre, il permet à l'homme d'appréhender son environnement par la perception tactilo-kinesthésique, de saisir et d'agir sur les objets.

Dans ce chapitre, nous nous pencherons sur deux directions de recherche distinctes, les interfaces haptiques et les interfaces tangibles. Nous verrons en quoi elles modifient les modalités de la relation homme-machine en ouvrant la voie à de nouveaux paradigmes d'interaction. Ces interfaces font l'objet d'une attention toute particulière dans le milieu de la recherche mais également dans le milieu industriel et muséographique et tendent à se démocratiser auprès des particuliers. D'une part, nous verrons que ces interfaces déploient chacune des moyens différents d'interagir avec un système interactif et qu'elles redéfinissent ainsi la notion d'instrument de manipulation telle qu'elle est proposée par le paradigme WIMP, notamment en délaissant la souris au profit de nouveaux périphériques d'entrée utilisateur.

D'autre part, nous verrons à travers ces modalités d'interaction qu'elles tendent toutes à considérer le geste comme un canal de communication privilégié dans la relation homme-machine. Nous tenterons de révéler ici les éléments physiques et conceptuels nécessaires à la mise en œuvre de tels systèmes. L'étude des interfaces haptiques nous donnera des éléments de compréhension sur la création d'illusions de contacts tactilo-kinesthésiques par l'intermédiaire d'instruments physiques. L'étude des interfaces utilisateur tangibles nous permettra de faire la distinction entre les interfaces tangibles et les interfaces tactiles, puis nous donnera des informations déterminantes pour comprendre en quoi ces interfaces s'inscrivent en des points différents du continuum de virtualité. Enfin, leur analyse nous donnera des éléments de compréhension des notions d'haptique, de feedback, de réalisme

²⁶ CADOZ Claude. *Le Geste canal de communication homme-machine : la communication instrumentale*. In TSI (Traitement du Signal et des Images), vol. 13. 1994, pp. 31-61 (coll. Hermès)

et d'instruments qui serviront, lors des prochains chapitres, à la mise en évidence de la faible granularité des interfaces multi-tactiles.

Terminologie générale

L'emploi indifférencié des notions d' « interaction » et d' « interactivité » dans les débats et les communications autour du sujet mérite que l'on revienne sur l'origine de ces deux termes pour en apprécier toutes les dimensions mais également pour s'affranchir des lieux communs et des définitions trop hâtives généralement associées à ces termes. « *Un dictionnaire commencerait à partir du moment où il ne donnerait plus le sens mais les besoins des mots.*²⁷ »

Le terme « interactivité » apparaît à la fin des années 1970 aux Etats-Unis puis quelques années plus tard dans le *Petit Robert* qui en donne la définition suivante : « *Activité de dialogue entre l'utilisateur d'un système informatique et la machine, par l'intermédiaire d'un écran.*²⁸ » Il est dérivé du mot « interactif » qui est alors relatif au domaine de l'informatique et défini par ce même dictionnaire par « *Qui permet d'utiliser le mode conversationnel.* » Depuis les années 1960 jusqu'à nos jours, nous utilisons le verbe *interagir* pour qualifier l'action réciproque entre deux éléments, humains ou non humains, pensant même aujourd'hui l'apparenter à « interactivité ». Or il semblerait que le terme « interagir », antérieur au terme « interactivité », soit directement issu de « interaction », terme beaucoup plus ancien qui relèverait de la définition suivante : « *Influence réciproque de deux phénomènes, de deux personnes.* » La confusion entre les termes « interaction » et « interactivité » semble donc liée à l'emploi inconsidéré d' « interagir » dans des champs disciplinaires comme l'informatique, la sociologie, la politique, les sciences de l'information et de la communication, etc. Aujourd'hui, tout interagit avec tout, sans que l'on sache bien ce que cela signifie.

Les définitions qui sont proposées dans les années 1990 reflètent ce que Fred Forest appelle l' « *esthétique de la communication* » : il conçoit l'interactivité en terme communicationnel comme l'influence des échanges de flux d'informations entre l'auteur et les spectateurs,

²⁷ BATAILLE Georges. *Informe*. In *Œuvres complètes I*, Denis Hollier. Paris, Gallimard, 1970, 464 p. (Coll. blanche)

²⁸ Le Petit Robert. *Dictionnaire alphabétique et analogique de la langue française*. Le Robert, (Coll. Dictionnaires Generalistes), 1996

mais également entre les spectateurs eux-mêmes, dans la naissance d'une coproduction de sens. « *L'interactivité est le degré d'implication auquel les participants d'un processus de communication contrôlent et échangent leurs rôles et leurs informations de façon réciproque*²⁹. » Les interactions concernent davantage les modalités des échanges entre les participants.

La dissection étymologique du terme « interactivité » par Douglas Edric Stanley³⁰ en 1997, nous éclaire davantage sur la capacité du mot à convenir au concept tel que je l'envisage. Il décompose ainsi « interactivité » en trois mots : in + ter + activité. « Activité » est un dérivé d'« acte » et participe passé passif substantivé du latin « *agere* » (faire). Or là où *facere* signifie également « faire » dans une relation à l'instant, *agere* représente l'activité dans une continuité. *Agere* viendrait également du grec *agein* (pousser devant soi) et ne concerne pas au départ la réalisation d'une tâche mais plutôt la « conduite » qui mène à son aboutissement. « Inter » se compose de « in » qui signifie en latin « en » et de « ter » qui sert à opposer deux éléments. « Inter » signifie donc littéralement « à l'intérieur de deux », dans le sens d'un espacement, de l'occupation d'un « entre », beaucoup plus que la dynamique de réciprocité telle que le terme « inter » a actuellement tendance à signifier. « *D'un côté on voit inter- en tant qu'espacement purement mesurable, compréhensible à l'espace même. Une logique de l'espace. De l'autre, on voit un espace d'opérations ou d'activations où l'espace lui-même est travaillé, de l'intérieur, par une force qui le dépasse.*³¹ » L'« interactivité » peut donc s'interpréter comme la maintenance, l'actualisation d'une dynamique entre deux éléments ou plus, dont l'activité est entretenue par les multiples actions entre ces éléments : les interactions.

²⁹ FOREST Fred. *Pour un art actuel - L'art à l'heure d'internet*. L'Harmattan. Paris, 1997, 282 p. (coll. L'Art en bref)

³⁰ STANLEY Douglas Edric. *Mode d'emploi : lexique de l'interactivité*. DEA Esthétique, Sciences et Technologies des Arts. Paris, Université de Paris VIII, 1997

³¹ *Ibid.*

1. Interaction homme-machine

1.1. Interaction homme-machine

De manière générale, nous pouvons décrire l'IHM (Interaction Homme-Machine ou Interface Homme-Machine) comme un domaine de recherche qui étudie les interactions entre l'homme et la machine. Cette discipline se situe à l'intersection de plusieurs domaines scientifiques comme l'informatique, les sciences cognitives, la psychologie et l'esthétique et couvre l'ensemble des études portant sur les interactions entre les utilisateurs et l'ordinateur par l'intermédiaire d'interfaces physiques (clavier, souris, etc.) et logicielles (interface graphique, univers virtuel, etc.)

L'ACM (*Association for Computing Machinery*, ou littéralement association pour la machinerie informatique) est un pôle d'intérêt en informatique fondé en 1947 dont la contribution scientifique a été l'organe de référence principal dans le milieu de la recherche et de l'enseignement entre 1960 et 1980. L'ACM définit les IHM comme « *la discipline qui s'intéresse au design, à l'évaluation et à l'implémentation de systèmes informatiques interactifs à l'usage de l'homme ainsi que l'étude des principaux phénomènes qui les entourent* ».³² » D'une part, cette acception reprend les deux définitions classiques de la science informatique définies par Newell, Perlis et Simon³³ en 1968 comme « *la science des ordinateurs et les principaux phénomènes qui les entourent* »³⁴ et par Denning, P. J., et al.³⁵ en 1988 comme « *l'étude rationnelle des processus algorithmiques qui définissent et transforment l'information : leur théorie, analyse, design, efficacité, implémentation, et application* »³⁶ ». D'autre part, la définition proposée par l'ACM caractérise la spécificité de l'IHM en évoquant les notions fondamentales de la relation homme-machine telles que

³² « *Human-computer interaction is a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them.* » In ACM SIGCHI. *Curricula for Human-Computer Interaction*. Rapport technique. ACM, New York, USA, 1992, disponible ici : http://old.sigchi.org/cdg/cdg2.html#2_1. Traduction libre : COQUELIN Marine et MORIN Rudy.

³³ NEWELL Allen, PERLIS A.J. et SIMON H.A. *What is computer science?* [Lettre à l'éditeur]. In *Science* 157, 1967, pp. 1373-1374

³⁴ *ibid.*

³⁵ DENNING Peter, COMER Douglas E., GRIES David, MULDER Michael C., TUCKER Allen B., TURNER A. Joe et YOUNG Paul R. *Computing as a discipline: preliminary report of the ACM task force on the core of computer science*. In *SIGCSE Bulletin* (20, 1), 1998, pp. 41-41

³⁶ *ibid.*

l'interactivité et la dimension « humaine » des interactions. Ils déclinent ainsi la recherche en IHM selon quatre axes :

- Les usages et le contexte liés aux ordinateurs,
- Les caractéristiques humaines,
- Les systèmes informatiques et le design d'interfaces,
- Les processus de développement.

Le design et l'évaluation de systèmes interactifs concernent des domaines industriels comme la communication, l'accès à l'information (internet par exemple), le domaine militaire, le jeu vidéo et bien d'autres contextes où les technologies et les utilisateurs sont dans un état constant de flux, d'évolution et de co-evolution. Les théories, les modèles et les systèmes conçus par les ingénieurs tentent de répondre à des besoins utilisateurs spécifiques en intégrant dans leurs approches les notions de satisfaction, d'efficacité et de réduction de marge d'erreurs. Les modèles d'interaction classiques de type WIMP (*Window, Icon, Menu, Pointing Device*) ou les modèles plus récents comme l'interaction instrumentale sont issus de travaux de recherche pluridisciplinaires qui ont progressivement déplacé le point focal des études en interaction homme-machine de l'ordinateur à l'homme. Les designers d'interaction sont continuellement confrontés à de nouveaux contextes, de nouvelles technologies d'affichage et d'acquisition, et à des scénarios d'usage différents qui les encouragent à adopter de nouveaux modèles d'interaction.

L'IHM est un domaine contraint à de multiples réactualisations tant la nature de son objet d'étude est polymorphe et soumis à certains facteurs sociaux, économiques et technologiques. La diminution des coûts, la miniaturisation des ordinateurs, l'augmentation permanente des vitesses de calcul, l'évolution des technologies d'affichage et d'interactions, et l'apparition de nouveaux « objets interactifs » créent des conditions d'utilisation et d'interactions qui suscitent de nouveaux enjeux et soulèvent de nouvelles problématiques. Concevoir l'IHM comme un champ d'étude visant à améliorer les interactions entre les utilisateurs et l'ordinateur suppose de s'intéresser à l'ensemble des modalités, des processus et des représentations de la relation homme-machine et cela implique le recours à la psychologie, aux sciences de l'information et de la communication, à la sociologie, à l'anthropologie, etc.

L'objectif de l'IHM est de mettre au point ou d'améliorer l'utilisation, l'efficacité et l'usabilité des systèmes informatiques. L'organisation internationale de normalisation (ISO) définit l'usabilité comme « *le degré selon lequel un produit peut être utilisé, par des utilisateurs identifiés, pour atteindre des buts définis avec efficacité, efficience et satisfaction, dans un contexte d'utilisation spécifié.* » Selon Jakob Nielson³⁷ et Ben Shneiderman³⁸, l'usabilité (ou utilisabilité) est généralement associée aux caractéristiques suivantes :

- *L'apprentissage* : la facilité pour un utilisateur de réaliser une tâche pour la première fois.
- *L'efficience* : la rapidité d'exécution des tâches après apprentissage.
- *La mémoire* : la facilité de la reprise en main du système après un moment sans utilisation.
- *Les erreurs* : le nombre d'erreurs commises par l'utilisateur, leur gravité et leur correction.
- *La satisfaction* : le degré de satisfaction de l'utilisation de l'interface par l'utilisateur.

L'IHM étudie les facteurs humains et technologiques impliqués dans la relation homme-machine afin de créer ou d'améliorer l'usabilité des interfaces.

1.2. Interface WIMP

Le début des années soixante marquent l'histoire des IHM puisque c'est une période incroyablement fertile où se cristallisent des concepts majeurs comme l'environnement graphique, la manipulation directe et l'ordinateur personnel. C'est un moment important de l'histoire des sciences informatiques, marqué par des avancées conceptuelles et techniques décisives qui ont défini les principes d'interaction homme-machine toujours utilisés aujourd'hui dans la majorité des programmes informatiques.

Parmi les chercheurs qui ont apporté une contribution scientifique capitale au domaine des IHM, Douglas Engelbart fait figure de pionnier. Dès 1959, il travaille sur un environnement de développement (*framework*) conceptuel, dont les enjeux entrent déjà en résonance avec la définition moderne de l'IHM. Ce travail de recherche s'inscrit dans le cadre d'un

³⁷ NIELSEN Jakob. *CHI'88 Trip Report*. Washington D.C., 15-19 mai 1988

³⁸ SHNEIDERMAN Ben. *Software Psychology : human factors in computer and information systems*. In *Winthrop Computer Systems Series*. Winthrop Publishers, 1980

contrat avec le bureau de recherche scientifique des sciences de l'information de l'US Air Force et aboutit en 1962 au rapport³⁹ intitulé « *Augmenter l'intellect humain : un environnement de développement conceptuel* ». Conscient de l'intensification des problèmes d'organisation sociale liés à la croissance et la complexification des systèmes de communication et d'information, il fait le constat que la technologie évolue dans le sens de l'automatisation des tâches et qu'elle n'a pas intégré dans son évolution les notions sociales comme les médias, le langage, le savoir et les compétences. Il voit dans l'outil informatique un moyen d'augmenter l' « intellect humain », terme qu'il associera plus tard à celui d' « intelligence collective » en orientant l'ordinateur vers un outil destiné à faciliter et accélérer les tâches essentielles liées à l'organisation et au savoir de la société.

« Par augmenter l'intellect humain, nous entendons augmenter la capacité de l'homme à appréhender des situations de problèmes complexes, mieux comprendre ses besoins particuliers, et trouver des solutions aux problèmes.⁴⁰ »

Par la suite Douglas Engelbart fonde l'ARC (*Augmentation Research Center*), un centre de recherche où vont naître des inventions révolutionnaires comme la souris, le premier système hypertexte fonctionnel, le multifenêtrage et des concepts comme les notions d'efficacité, de réduction d'erreurs, de fatigue musculaire et de stress, qui figurent parmi les enjeux principaux de la recherche en IHM.

Le concept de manipulation directe, où les objets graphiques de l'interface sont directement manipulables avec la souris, a été présenté pour la première fois par Ivan Sutherland⁴¹ sur le *Sketchpad*, programme informatique qu'il réalisa dans le cadre de sa thèse en 1963. Cette application qui reçut par ailleurs le *Turing Award* en 1988 constitua un modèle pour l'évolution des interactions homme-machine. Les concepts modernes d'interface utilisateur graphique (*Graphical User Interface* ou GUI) et de programmation orientée objet se sont d'ailleurs très largement inspirés des travaux de Ivan Sutherland. Le programme *SketchPad*

³⁹ ENGELBART Douglas C. *Augmenting human intellect : a conceptual framework*. Report. Air force office of scientific research, Washington 25, C.C., 1962, 133 p.

⁴⁰ « *By augmenting human intellect we mean increasing the capability of a man to approach a complex problem situation, to gain comprehension to suit his particular needs, and to derive solutions to problems.* » In *ibid*. Traduction libre : COQUELIN Marine et MORIN Rudy.

⁴¹ SUTHERLAND Ivan Edward. *SketchPad : A man-machine graphical communication System*. In *AFIPS Spring Joint Computer Conference*. 1963, pp. 329-346

supportait la manipulation d'objets graphiques par l'intermédiaire d'un crayon optique, comme saisir un objet, le déplacer ou modifier sa taille. David Canfield Smith, designer d'interfaces pour Xerox, prolonge les travaux d'Ivan Sutherland dans sa thèse sur un programme informatique appelé Pygmalion⁴² en proposant le terme « icône » pour nommer la représentation graphique d'entités ou de processus informatiques abstraits. Le terme de manipulation, que nous rapprocherons du latin populaire qui signifie « conduire par la main », désigne ici la modification des propriétés d'un objet graphique, tant ses propriétés spatiales que ses propriétés intrinsèques (forme, couleurs, géométrie, etc.) par l'action d'un périphérique d'entrée.

Dans les années soixante-dix, Xerox fonde le PARC à Palo Alto, un laboratoire de recherche qui sera à l'origine de nombreuses avancées techniques et conceptuelles dans les domaines des interfaces graphiques et de la manipulation directe. Alan Kay, chercheur au sein des laboratoires PARC et considéré comme un des pères de la programmation orientée objet, démocratise le concept d'interfaces à manipulation directe dans un article⁴³ sur le *Dynabook*, idée née de la volonté de créer un ordinateur à destination des enfants, une sorte de tablette PC aux vertus éducatives. Conscient de l'avant-gardisme de son projet et des contraintes techniques liées à la miniaturisation d'un tel dispositif, Alan Kay commence son étude en indiquant le caractère fictionnel du projet qu'il présente :

« Bien que cette contribution puisse être lue comme de la science-fiction, certaines tendances en matière de miniaturisation et de réduction des coûts garantissent quasiment que bien des notions présentées se concrétiseront dans un futur proche.⁴⁴ »

Alan Kay s'inspire en particulier des contributions scientifiques et des théories amenées par les sciences cognitives et la psychologie. Jean Piaget et Jerome Bruner ont fait l'observation que la compréhension de notre environnement est facilitée par la stimulation visuelle, l'expérience et la manipulation des objets qui le composent. « *On ne connaît un objet qu'en*

⁴² SMITH D.C. *Pygmalion : A computer program to model and stimulate creative thought*. Thèse. Basel, Stuttgart: Birkhauser Verlag, Stanford University Computer Science Department, 1975

⁴³ KAY Alan. *A personal computer for children of all ages*. In *Proceedings of the ACM national conference*, Boston Aug. Xerox Palo Alto Research Center, 1972

⁴⁴ « *Although it should be read as science fiction, current trends in miniaturization and price reduction almost guarantee that many of the notions discussed will actually happen in the near future.* » In *ibid*. Traduction libre : COQUELIN Marine et MORIN Rudy.

agissant sur lui et en le transformant.⁴⁵ » Deux notions fondamentales de Jean Piaget se retrouvent dans les analyses d'Alan Kay. La première est que le savoir, en particulier celui des jeunes enfants, est restreint à une série de modèles opératoires, chacun est en quelque sorte *ad hoc* et n'a pas besoin d'être cohérent avec les autres. Ce n'est que plus tard au cours du développement que la logique intervient et cela même à travers des situations paradoxales. La seconde notion est que le développement s'effectue selon une succession d'étapes qui semblent indépendantes de l'environnement culturel : le stade de l'intelligence sensori-motrice, le stade de l'intelligence préopératoire, le stade des opérations concrètes et le stade des opérations formelles. L'approche constructiviste interactionniste proposée par Jerome Bruner participe également aux analyses d'Alan Kay en ce sens où elle définit l'apprentissage autour de l'idée d'un sujet actif construisant de nouvelles idées ou concepts à partir de connaissances existantes. Le processus de la construction du savoir se fait, selon Bruner, de manière active et implique la transformation de l'information, issue de l'expérience, de la création d'hypothèses et de la prise de décision. Jerome Bruner distingue trois étapes de la représentation : le mode éactif (mémoire de l'action physique et musculaire), le mode iconique (images mentales) et le mode symbolique (langage et symboles).

En 1973, Alan Kay et Chuck Thacker développe un prototype du *Dynabook*, nommé *Alto*, ordinateur personnel principalement développé autour du concept central amené par Douglas Engelbart, à savoir la notion de manipulation directe, et en tenant compte des modèles de la représentation et de la construction du savoir proposés par Jean Piaget et Jerome Bruner. Ce prototype, composé d'un couple clavier/souris en entrée, d'un écran pour la visualisation, dispose de la toute première interface graphique exploitée par Xerox. Elle se compose de fenêtres et utilise des métaphores visuelles et fonctionnelles comme des dossiers pour les documents ou un curseur pour montrer les déplacements de la souris. Mergouza Wilberts, en 1980, donne un nom pour ce type d'interactions qu'il appelle WIMP pour *Window, Icon, Menu, Pointing device* (fenêtre, icône, menu, dispositif de pointage). Alan Kay conçoit l'ordinateur comme un système abstrait et complexe qu'il faut organiser pour faciliter son utilisation et ceci sans avoir à se rappeler de commandes compliquées. En simulant l'aspect et le comportement familier de l'environnement par l'intermédiaire d'objets graphiques comme les documents, les outils ou les menus déroulants, les fonctions

⁴⁵ PIAGET Jean. *Psychologie et épistémologie*, Paris, Denoël, 1970, 85 p.

de l'ordinateur sont rendues visibles et représentées dans un contexte logique et prédictif. Les idées et les concepts mis en avant par Alan Kay sont fondamentaux pour saisir le changement paradigmatique amené par ce type d'interfaces.

Ce n'est qu'en 1983, que Ben Shneidermann⁴⁶ nommera ce type d'interactions « *manipulation directe* », par opposition à l'utilisation quasi systématique de langages de commandes (*Command Line Interface* ou CLI) jusqu'aux années 1980. Ce type d'interactions fait appel à trois techniques :

- Fournir un moyen physique et direct de déplacer un curseur ou de manipuler les objets d'intérêt,
- Offrir une représentation visuelle concrète des objets d'intérêt et afficher immédiatement la représentation des modifications apportées,
- Éviter d'utiliser un langage informatique et prendre en compte le modèle cognitif utilisé pour concevoir l'affichage.

Les premières exploitations commerciales de systèmes informatiques proposant des interfaces à manipulation directe ont été successivement le *Xerox Star*⁴⁷ en 1981, le *Apple Lisa*⁴⁸ en 1982 et le *Macintosh*⁴⁹ en 1984. Depuis les années 1980 et encore aujourd'hui, la grande majorité des interactions homme-machine fonctionnent sur les standards créés à partir des travaux et des idées avant-gardistes de Douglas Engelbart, Ivan Sutherland et Alan Kay.

La généralisation des environnements graphiques peut s'expliquer par les nombreux avantages qu'ils apportent. Les tâches utilisateur se trouvent grandement simplifiées car la représentation graphique des objets les rend accessibles plus rapidement. La matérialisation d'outils sous la forme d'icônes, facilement mémorisables, rend l'utilisation de l'ordinateur plus simple et intuitive. Technique implémentée depuis l'interface utilisateur *Star* de *Xerox*, la métaphore du « bureau » de travail physique est clairement représentée et incarnée dans

⁴⁶ SHNEIDERMAN Ben. *Direct Manipulation : a step beyond programming languages*. In *Computer* (16), 1983, pp. 57-69

⁴⁷ SMITH David Canfield, IRBY Charles, KIMBALL Ralph et HARSLEM Eric. *The star user interface: an overview*. In *AFIPS Conference Proceedings*, vol. 55. National Computer Conference, AFIPS Press, Arlington, VA, USA, 1986, pp. 383-396

⁴⁸ WILLIAMS Gregg. *The Lisa Computer System*. In *Byte Magazine*. 1983, pp. 33-50

⁴⁹ WILLIAMS Gregg. *The Apple Macintosh Computer*. In *Byte Magazine*. 1984, pp. 30-54

les icônes visibles de l'interface. Les icônes sont déplaçables et peuvent être réarrangées afin de correspondre aux conditions de travail optimales pour l'utilisateur, un peu comme l'on rangerait ses affaires sur un bureau ordinaire. L'intention est de créer les moyens susceptibles de rendre l'utilisation de l'ordinateur plus « intuitive » et de diminuer la charge cognitive nécessaire à l'exécution des tâches. En sélectionnant les objets d'intérêt et en les modifiant, l'utilisateur peut observer les changements instantanément : c'est l'apparition de la notion de feedback, ou rétroaction, dans les interfaces interactives. Enfin, la manipulation directe est accessible aux novices qui s'habituent à l'outil informatique et peuvent progresser sans pour autant en avoir un usage régulier.

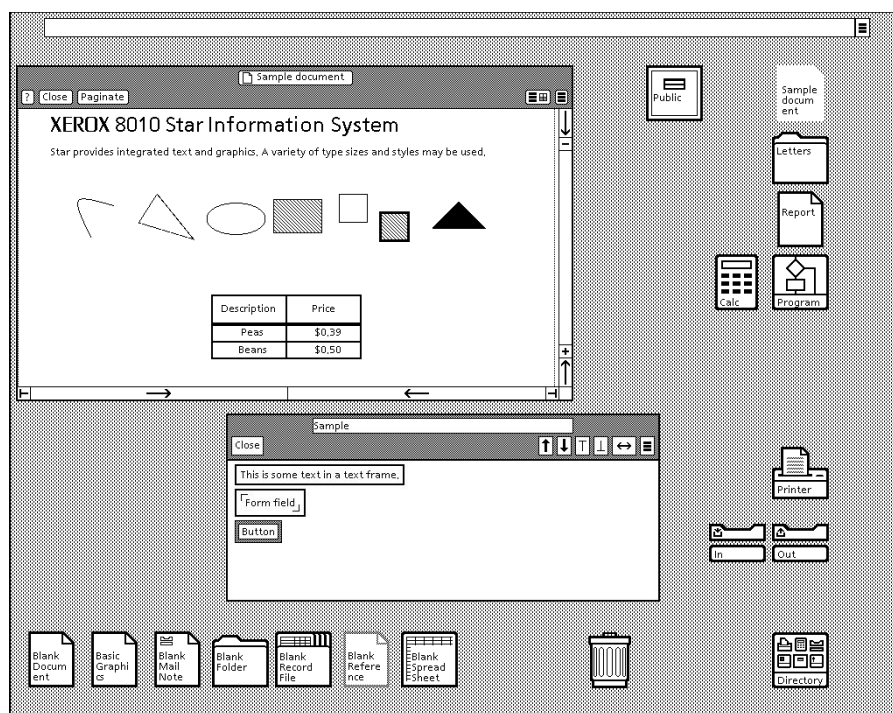


Figure 1 - Interface graphique du Xerox Star reprenant les principes déjà présents dans l'ordinateur Alto, 1981.

1.3. Limites du modèle WIMP

En 1997, Michel Beaudouin-Lafon mène une étude portant sur dix applications très largement répandues sur *Apple Macintosh* et qui concernent des domaines aussi variés que le traitement de texte, la retouche d'image et la modélisation 3D. En partant des modèles d'évaluation précédents, il applique une grille d'analyse quantitative et qualitative afin de mesurer la qualité de l'interface, des outils et des interactions de ces logiciels. Il évalue

également le degré de correspondance entre les modèles d'interaction utilisés et les principes de manipulation directe définis par Ben Shneiderman⁵⁰ en 1983 :

- *La représentation continue des objets d'intérêt.* Les résultats montrent que les applications étudiées ne respectent pas forcément les principes de la manipulation directe, ou plus précisément qu'elles ne les interprètent pas correctement. L'accessibilité des objets d'intérêt, qui doivent être présents à tout moment, est souvent visuellement mise en retrait à cause du nombre croissant d'outils et d'objets graphiques qui n'appartiennent pas au domaine des objets d'intérêt. Les boîtes de dialogues, feuilles de style ou calques graphiques prennent une part importante dans le modèle mental de l'utilisateur qui les perçoit comme objets d'intérêt alors qu'ils ne sont que rarement implémentés comme tels.
- *L'action physique sur les objets plutôt qu'une syntaxe complexe.* Les périphériques d'entrée habituels sont le clavier et la souris. En considérant d'une part le nombre limité d'actions disponibles (taper du texte, utiliser des raccourcis clavier, pointer, cliquer et déplacer) et d'autre part l'étendue des commandes et des actions proposées par le programme, les interfaces font obligatoirement appel à l'utilisation de menus et de boîtes de dialogues qui nécessitent des étapes supplémentaires avant d'obtenir l'effet désiré sur l'objet d'intérêt. Le concept de manipulation directe est mis à mal en ce sens où la manipulation est soumise à une série de commandes et les interactions sont le plus souvent indirectes et pas forcément plus efficaces que les interfaces en ligne de commande.
- *Des opérations rapides, incrémentales et réversibles avec un effet immédiat visible sur les objets d'intérêt.* Les opérations ne sont pas rapides car l'utilisation des boîtes de dialogue implique le temps de la sélection, de la saisie d'un champ texte et de la validation par un bouton « ok » ou « valider ». Les opérations ne sont donc pas toutes incrémentales car si le résultat sur l'objet d'intérêt n'est pas satisfaisant après la validation de la boîte de dialogue, il faut donc recommencer toutes les étapes de la modification. Ceci est particulièrement dommageable lorsque l'utilisateur travaille

⁵⁰ SHNEIDERMAN Ben. *Direct Manipulation : a step beyond programming languages*. In Computer (16), 1983, pp. 57-69

par la méthode d'essai / erreur, c'est à dire lorsqu'il teste toute une série de valeurs jusqu'à l'obtention d'un résultat satisfaisant.

- *Une approche en couches ou en spirale qui permet d'utiliser le programme avec un minimum de connaissances.* Bien que les interfaces WIMP proposent des interactions plus simples à utiliser et à apprendre, l'utilisation de combinaisons de touches ou l'association clavier/souris pour l'exécution d'une commande n'est pas toujours la même d'un programme à un autre, d'une version à une autre, et cela rend plus difficile la mémorisation et l'apprentissage.

La difficulté de mettre en application les principes de la manipulation directe dans les interfaces WIMP s'explique en partie parce que les programmes sont de plus en plus complexes et les nombreuses commandes sont accessibles par l'intermédiaire de boîtes de dialogues qui sont plus explicites pour l'utilisateur mais cependant moins efficaces. De plus, le couple clavier/souris comme périphérique d'entrée utilisateur est déjà largement répandu en 1983 et cela crée des conditions d'utilisation qui imposent aux designers d'interaction de simplifier l'interface et de réduire la charge cognitive nécessaire aux interactions utilisateur.

2. Instrument d'interaction

2.1. Périphérique d'entrée

Jusque dans les années 1990, le design et la fabrication des périphériques utilisateur répondaient davantage à des contraintes de faisabilité technique qu'à une réelle volonté d'améliorer l'efficacité de ces dispositifs, en tenant compte des facteurs de performance utilisateur. Selon Robert J.K. Jacob *et al.*⁵¹, l'efficacité d'un périphérique d'entrée pour une tâche déterminée dépend, au-delà de la structure physique du dispositif, de la structure perceptuelle de la tâche à réaliser, du périphérique et des interrelations entre la structure perceptuelle de la tâche et les paramètres de contrôle du périphérique.

⁵¹ JACOB Robert J. K., SIBERT Linda E., McFARLANE Daniel C. et MULLEN M. Preston Jr. *Integrity and separability of input devices*. In *ACM Trans. Computer Human Interaction* (1, 1), 1994, pp. 3-26

Dans les années 1980, les périphériques d'entrée utilisateur étaient divisés en classes logiques selon leur capacité à réaliser les opérations utilisateur élémentaires. En 1984, James Foley *et al.*⁵² proposent une taxinomie des modèles d'interaction qui suppose que plusieurs techniques d'interaction sont susceptibles de réaliser une tâche élémentaire avec la même efficacité, tant que les dispositifs sont capables d'envoyer les mêmes informations au système. Cette taxinomie repose sur la distinction de six tâches d'interaction génériques pour lesquelles ils associent des dispositifs physiques capables de réaliser la tâche. Dès 1986, William Buxton⁵³ souligne l'importance des propriétés physiques des dispositifs et leurs implications dans la perception de l'utilisateur. Il propose donc un système de classification dont le niveau pragmatique se décline selon deux axes : le nombre de dimensions et le type de propriétés captées. En 1991, Stuart K. Card *et al.*, devant le développement croissant de nouveaux périphériques d'entrées : souris, stylet, tablettes graphiques, gants et capteurs, proposent une série de méthodes afin d'analyser et de systématiser les périphériques d'entrée à travers l'analyse morphologique du design d'interaction⁵⁴. Il s'agit essentiellement de définir des moyens concrets pour analyser, évaluer et classer les périphériques d'entrée existants ou à venir. En complétant les taxinomies proposées par Buxton et Baecker^{55 56}, fondées sur les propriétés physiques et le nombre de dimensions spatiales du périphérique, et en pointant la nécessité d'inclure la totalité des facteurs liés à son utilisation, Card *et al.* proposent un espace de conception (*design space*) permettant d'étudier la performance des entrées utilisateur en se basant sur les principes suivants :

- L' « empreinte sur le bureau ». La quantité d'espace utilisée par le périphérique sur le bureau,
- La vitesse du pointeur. Combien de temps est nécessaire au dispositif pour sélectionner une cible,

⁵² FOLEY James D., WALLACE Victor L. et CHAN Peggy. *The human factors of computer graphics interaction techniques*. In *IEEE Computer Graphics and Applications*, 1984, pp. 13-48

⁵³ BUXTON William. *There's more to interaction than meets the eye: Some issues in manual input*. In *User Centered System Design: New Perspectives on Human-Computer Interaction*. Lawrence Erlbaum, Hillsdale, 1986, pp. 319-337

⁵⁴ CARD Stuart K., MACKINLAY Jock D. et ROBERTSON George G. *A morphological analysis of the design space of input devices*. ACM, 1991, pp. 99-122

⁵⁵ BAECKER R. M. et BUXTON William. *EDS. Readings in Human-Computer Interaction: A Multidisciplinary Approach*. Kaufmann, Los Altos, Calif., 1987, pp. 357-365

⁵⁶ BUXTON William. *Lexical and pragmatic considerations of input structures*. In *SIGGRAPH Computer Graphics*. 17, 1 1983, pp. 31-37

- *La précision du pointeur.* La plus petite taille pour un objet cible facilement sélectionnable,
- *Les erreurs.* Le pourcentage de réussite de la sélection, qualité de la ligne dessinée, etc.,
- *Le temps d'apprentissage.* Le temps nécessaire avant que l'utilisateur utilise le périphérique efficacement,
- *Le temps pour saisir le périphérique.* Le temps qu'il faut pour attraper le périphérique si les mains sont occupées à autre chose,
- *Les préférences utilisateur.* A quel point l'utilisateur apprécie le périphérique et à quel point il le préfère aux autres,
- *Le coût.* Le prix de vente du périphérique.

Le rôle principal d'un périphérique d'entrée est de transmettre l'information et les intentions de l'utilisateur à l'ordinateur. Afin d'améliorer l'efficacité de telles interfaces physiques, des progrès conséquents ont été fait sur la vitesse de transmission des données, sur le nombre de degrés de liberté et sur l'ergonomie du dispositif d'interaction.

D'un côté du canal de communication homme-machine, la nature des interactions est déterminée par les propriétés physiologiques de l'homme et de l'autre par les modalités d'interaction des périphériques d'entrée. Au cours des années 1990, les recherches menées sur les périphériques d'entrée utilisateur ont permis de définir les axes de développement pour les interfaces homme-machine à venir. D'une part, bien que la souris soit restée le périphérique le plus répandu, notamment car les utilisateurs ont su faire preuve d'adaptabilité, il apparaît très tôt que l'efficacité de tels dispositifs est dépendante du type de tâche à réaliser et dès lors, les efforts de la recherche se sont concentrés sur le développement et l'évaluation de nouveaux périphériques. D'autre part, la recherche en IHM s'est orientée vers de nouveaux canaux de communication homme-machine en se concentrant davantage sur les modes de la perception humaine (sens, organes) et sur les processus cognitifs (mémoire, apprentissage) engagés dans la relation homme-machine.

2.2. Interaction instrumentale

L'introduction de menus, de boîtes de dialogue et de barres de défilement dans les interfaces fait clairement apparaître la notion d'instrument décrite par Michel Beaudouin-

Lafon. L'utilisation d'un outil comme médiateur unique entre l'utilisateur et l'objet d'intérêt ne suit pas forcément les principes de la manipulation directe mais il correspond davantage à la manipulation d'objets telle qu'on en a l'habitude dans le monde physique. En effet, nous utilisons des pinceaux pour peindre, un interrupteur pour allumer ou éteindre la lumière ou bien une loupe pour rendre un texte plus lisible. La manipulation instrumentale est basée sur la manière dont nous utilisons « naturellement » les outils ou les instruments pour manipuler des objets dans le monde réel.

L'interaction instrumentale peut être considérée comme une extension de la manipulation directe car elle généralise d'une part les concepts d'interaction proposés par les interfaces de type WIMP et tente de donner un cadre de compréhension pour les modèles d'interactions émergents comme la réalité augmentée, les interfaces tangibles et multi-modales ou la manipulation bi-manuelle. L'interaction entre l'utilisateur et le domaine des objets (interface) s'effectue par l'intermédiaire d'instruments interactifs, comparables à des outils ou des instruments tels qu'on les utilise dans le monde physique. Selon Michel Beaudouin-Lafon⁵⁷, les principales motivations du modèle de l'interaction instrumentale sont :

- Généraliser les techniques d'interaction amenées précédemment par les travaux réalisés dans les années 1980 sur la manipulation directe et les concepts liés aux interfaces graphiques utilisateur,
- Proposer des méthodes qualitatives et quantitatives pour l'analyse des modèles d'interaction afin de faciliter les choix des designers d'interface confrontés à un problème particulier,
- Définir un cadre d'analyse pour identifier les voies encore inexploitées de l'IHM et qui pourrait ouvrir sur de nouvelles techniques d'interaction,
- Proposer des outils de développement qui permettent d'intégrer plus simplement de nouvelles techniques d'interaction dans les programmes informatiques.

⁵⁷ BEAUDOUIN-LAFON Michel. *Instrumental interaction : an interaction model for designing post-WIMP user interfaces*. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (CHI '00). ACM, New York, USA, pp. 446-453

Le terme « instrument », dans sa forme classique, désigne le médiateur entre l'utilisateur et l'objet d'intérêt de l'application et renvoie à la définition donnée par le *Petit Robert* : « *Objet fabriqué servant à exécuter quelque chose, à faire une opération (Instrument est plus général et moins concret que outil; désigne des objets plus simples que appareil, machine).*⁵⁸ » Michel Beaudouin-Lafon décrit l'instrument d'interaction comme un instrument qui opère sur des objets du domaine de l'application, en modifiant ses attributs et propriétés, de manière explicite (faire défiler une image en cliquant et déplaçant la barre de défilement) ou implicite (utilisation de la molette de la souris pour cette même action). Les retours d'information sont sur trois niveaux : sur l'instrument physique (notamment les dispositifs à retour de force), sur la représentation graphique de l'instrument et sur l'objet d'intérêt.

« Un instrument est composé d'une partie physique et d'une partie logique. La partie physique comprend les transducteurs d'entrée-sortie utilisés par l'instrument, en entrée pour capter l'action physique de l'utilisateur et en sortie pour lui présenter un retour d'information. [...] La partie logique de l'instrument comprend en entrée la méthode de transformation des actions de l'utilisateur sur l'instrument logique et en sortie la représentation de l'instrument. »⁵⁹

La manipulation d'objets graphiques dans une interface implique l'utilisation de périphériques d'entrée utilisateur et la modification des propriétés de l'objet dans l'interface graphique. La notion d'instrument ne fait pas exclusivement référence à l'outil physique et son analogon numérique, mais elle renvoie aussi à des concepts plus larges, réunificateurs, comme les styles de texte, les matériaux ou la palette d'outils. Ces concepts sont matérialisés à l'écran sous la forme d'objets graphiques, c'est le principe de réification avancé par Michel Beaudouin-Lafon. Ces instruments sont en quelque sorte des méta-instruments qui regroupent un ensemble d'outils plus spécifiques. Un instrument d'interaction est bien souvent la réunion de plusieurs commandes qui sont exécutées de manière transparente pour l'utilisateur.

⁵⁸ Le Petit Robert. *Dictionnaire alphabétique et analogique de la langue française*. Le Robert, (Coll. Dictionnaires Generalistes), 1996

⁵⁹ BEAUDOUIN-LAFON Michel. *Interaction Instrumentale : de la manipulation directe à la réalité augmentée*. IHM'97, 9èmes Journées sur l'Ingénierie de l'Interaction Homme-Machine. Cépadués Editions, 1997, pp. 97-104

Le modèle d'interaction instrumentale permet également d'appréhender des formes d'interaction bi-manuelles où la main non-dominante n'utilise pas forcément les touches de raccourcis du clavier en combinaison avec la souris mais d'autres périphériques de placement. C'est le cas des *Toolglass*, concept d'interface marginale, resté dans les laboratoires et présenté par Bier *et al.*⁶⁰ en 1993, qui consiste en une palette d'outils transparente qui vient se placer entre l'application et le curseur. Pendant que la main dominante modifie la position du curseur avec une souris, l'autre main positionne la palette d'outils avec un trackball par exemple et peut modifier le type d'outil actif. En amenant la palette transparente vers l'objet d'intérêt, le regard ne se détourne plus et l'ensemble des mouvements des mains tend vers une seule et même tâche. Parmi les fonctionnalités proposées par les *Toolglass*, il est possible d'appliquer des filtres visuels, les *Magic Lenses*, afin de mettre en valeur certaines informations ou au contraire de masquer les données inutiles. Ces « lentilles magiques » sont des filtres transformant la représentation des données dont l'exemple le plus explicite est celui de la loupe qui agrandit la zone située en dessous. Il est également possible de superposer les *Toolglass* et les lentilles afin de créer des combinaisons d'outils simplifiant l'exécution de la tâche ou de la manipulation.

2.3. Instrument haptique

Les technologies haptiques concernent les dispositifs et les interfaces qui, contrairement aux périphériques classiques (clavier/souris), intègrent dans la relation homme-machine une approche plus « naturelle » de l'utilisation des instruments dans la manipulation d'objets graphiques. L'instrument de la manipulation est matérialisé et externalisé sous la forme d'un périphérique d'entrée utilisateur qui va simuler l'utilisation de l'outil comme si l'utilisateur agissait dans le monde physique des objets. Les périphériques et interfaces haptiques ajoutent aux canaux visuels et audio la perception tactile et kinesthésique. Ils sont en effet capables de percevoir les informations mécaniques de la manipulation et de simuler des informations tactilo-kinesthésiques comme le retour de force. Les domaines d'exploitation sont multiples et concernent par exemple l'informatique, la médecine, la psychologie, la

⁶⁰ BIER Eric A., STONE Maureen C., PIER Ken, BUXTON William et DEROSE Tony D. *Toolglass and magic lenses: the see-through interface*. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, (SIGGRAPH '93). ACM, New York, USA, 1993, pp. 73-80

biologie et le jeu vidéo. Minsky *et al.* décrivent les technologies haptiques comme des dispositifs qui permettent de ressentir les objets virtuels :

« Les technologies à retour de force fonctionnent en utilisant des actuateurs mécaniques qui appliquent des forces sur l'utilisateur. En simulant la physique du monde virtuel, nous pouvons calculer ces forces en temps-réel, et les envoyer aux actuateurs afin que l'utilisateur les ressente. Ce que cela signifie réellement, c'est qu'une personne utilisant un système haptique peut ressentir la simulation d'un objet solide comme s'il se trouvait réellement face à lui.⁶¹ »

Terminologie

Au sens strict, l'haptique est un terme désignant la science qui étudie le toucher (haptique provient du grec *aptomai* ou *haptesthai* qui signifie « je touche » ou « toucher ») et les phénomènes de perception kinesthésique qui concernent la sensation de mouvement des parties du corps. Selon Hatwell *et al.*⁶², il existe deux types de perception tactile manuelle : la perception cutanée et la perception tactilo-kinesthésique. La perception cutanée, c'est la sensation des textures que nous éprouvons lorsque nous touchons un objet. Ce sont essentiellement les capteurs de pressions que nous avons sous la peau qui nous donnent des informations de relief sur les objets que nous saisissons. La perception tactilo-kinesthésique, terme introduit par György E. Révész⁶³, concerne la stimulation de la peau résultant des mouvements d'exploration de la main afin d'apprécier la forme d'un objet. Les processus complexes de l'exploration spatiale impliquent également les informations proprioceptives et motrices de la perception. Cela concerne la posture du corps, sa position, la vitesse et les forces physiques générées par les muscles.

Réalisme et simulation physique

La représentation de l'espace de travail virtuel par un bureau et celle des outils de la manipulation par des icônes étaient la première étape de la transition des interactions

⁶¹ MINSKY Margaret, MING Ouh-young, STEELE Oliver, BROOKS Frederick P., Jr. et BEHENSKY Max. *Feeling and seeing: issues in force display*. In *Proceedings of the 1990 symposium on Interactive 3D graphics (I3D '90)*. ACM, New York, USA, 1990, pp. 235-241

⁶² HATWELL Yvette, STRERI A. et GENTAZ E. *Toucher pour connaître. Psychologie cognitive de la perception tactile manuelle*. Paris, PUF, 2000, 332 p.

⁶³ RÉVÉSZ Géza. *System der optischen und haptischen Raumtäuschungen*. In *Zeitschrift für Physiologie*, 131, 1934, pp. 296-375

homme-machines vers une forme plus « naturelle », en profitant de la valeur symbolique des objets ou des instruments représentés à l'écran. Ces représentations graphiques numériques sont dites « virtuelles ». Elles sont la manifestation de transferts d'énergies électriques qui sont analysées et interprétées par les algorithmes informatiques. Avec les interfaces haptiques, ce qui change, ce sont les moyens qui permettent de simuler l'existence de ces objets en augmentant le réalisme de leur représentation et celui de leur manipulation. Claude Cadoz⁶⁴ parle alors de « *représentation intégrale* ». La perception visuelle et l'ouïe ne sont pas les seuls sens de l'homme impliqués dans la connaissance de son environnement physique. Les systèmes d'interaction classiques (clavier/souris) ne fournissent pas suffisamment d'informations au cours de l'exploration et de la manipulation d'objets virtuels pour que l'utilisateur ait la conviction de leur matérialité. James Jerome Gibson⁶⁵ définit les systèmes haptiques comme des instruments dont la fin est de donner à la relation homme-machine les caractères de la relation entre l'homme et son environnement, ceux de l'expérience du monde par l'exploration du corps.

Comparativement à la souris, qui ne possède que deux degrés de liberté, les technologies haptiques disposent généralement de trois à six degrés de liberté. Elles sont généralement composées de deux « organes » :

- Un système de captation chargé de convertir les phénomènes gestuels en signaux recevables par l'ordinateur,
- Un effecteur chargé de transformer les informations et les paramètres de l'objet virtuel en énergie physique (forces exercées sur l'utilisateur).

Le réalisme et la qualité des systèmes haptiques sont largement dépendants du taux de rafraîchissement (généralement un kHz ou plus) de la boucle action/perception des capteurs d'informations kinesthésiques, de la mise à jour de la représentation graphique et des effecteurs. Un autre facteur décisif est le degré de correspondance des forces émises en retour par le ou les transducteurs (capteurs et actuateurs). Il s'agit de donner l'illusion de la matérialité de l'objet virtuel en simulant ses propriétés physiques comme sa rigidité, sa texture, sa forme et parfois même sa température.

⁶⁴ CADOZ Claude. *Les réalités virtuelles*. Paris, Flammarion, 1994, Collection Dominos, 125 p.

⁶⁵ GIBSON James. J. *The Senses Considered as Perceptual Systems*. Greenwood Press Reprint, 1983, 335 p.

Simuler la manipulation d'un objet par l'intermédiaire d'un instrument d'interaction nécessite de saisir leur nature physique. Il existe deux types d'objets, animés (objets mécaniques, élastiques, etc.) ou inanimés (une feuille de papier sur un bureau). Les objets inanimés dissipent l'énergie fournie et appliquent une force opposée à celle qui leur est appliquée. Les objets animés sont capables de fournir de l'énergie sous la forme de forces localisées dépendant de leur nature (articulation, sources d'énergie, etc.) En fonction du type de manipulation à effectuer, nécessitant une précision et un réalisme plus ou moins importants, deux types de dispositifs haptiques existent : passifs et actifs. Les systèmes passifs consistent en dispositifs à dissipation programmable en fonction du temps ou en fonction de leur position. Les systèmes actifs sont des systèmes capables de restituer des échanges d'énergie plus complexes entre l'utilisateur et la machine dans le sens où ils sont en mesure de simuler plus efficacement les tensions exercées par l'objet ou les modifications des propriétés physiques de l'objet en temps réel. La conception de tels systèmes met en avant la notion d'échanges bidirectionnels, une fonctionnalité qui permet d'expliquer les sensations d'immédiateté et de réalisme ressenties par l'utilisateur.

Exemple 1 : le gant de données

Le gant de données (*dataglove*) est un gant muni de capteurs qui est utilisé afin de capturer les données physiques des mouvements de la main. Ce type de périphérique est largement utilisé dans les applications de réalité virtuelle car il permet de « virtualiser » la main de l'utilisateur en projetant un double numérique qui reproduit exactement les mêmes mouvements. En effet, ces gants peuvent être équipés de marqueurs magnétiques de localisation et d'accéléromètres qui renseignent sur la position et la rotation de la main tandis que d'autres capteurs indiquent le fléchissement des doigts donnant au système suffisamment d'informations pour interpréter l'ensemble des paramètres d'activation de la main. L'ensemble du dispositif fournit deux types d'information : des données de référencement global comme la position et la rotation dans l'espace réel et des données relatives comme l'angle que fait l'articulation entre deux segments de la main.

Le *DataGlove*, gant de données de VPL Research, a été présenté pour la première fois à la conférence CHI (conférence sur les interactions homme-machine) par Zimmerman *et al.*⁶⁶ en 1987.



Figure 2 - Le gant de données à fibres optiques conçu par Thomas Zimmerman en 1987, VPL Research.

Les performances de cette version étaient limitées et le *DataGlove* a rencontré de vives critiques concernant le manque de précision des capteurs et le nombre limité des positions de la main que le système pouvait détecter. Le *CyberGlove*, gant de données conçu par Virtual Technologies en 1990 constitua une solution plus fiable et plus confortable grâce à la modification du système de captation en fibre optique utilisé précédemment.

Les multiples possibilités offertes par le gant de données ont ouvert la voie à l'exploration de nouvelles techniques d'interaction et de manipulation d'objets en deux ou trois dimensions par l'intermédiaire de périphériques d'entrée à 3 degrés de liberté et plus. D'autres directions ont également été étudiées comme la navigation en temps réel dans un univers virtuel par des gestes déictiques ou le contrôle d'un système par des gestes de commande.

Exemple 2 : le bras haptique

Le bras à retour de force est un dispositif généralement composé d'un stylet monté sur un bras mécanique dont les jointures sont contrôlées par des servo-moteurs. Le stylet peut être remplacé par un autre instrument en fonction de la tâche à effectuer et il bénéficie de 6

⁶⁶ ZIMMERMAN Thomas G., LANIER Jaron, BLANCHARD Chuck, BRYSON Steve et HARVILL Young. *A hand gesture interface device*. SIGCHI Bull. 17, SI, 1986, pp. 189-192

degrés de liberté (x , y , z , rotation x , rotation y et rotation z). Ce périphérique donne la sensation de contact avec l'objet virtuel grâce à un retour de force programmable. Cela permet de donner à l'utilisateur des informations sur la texture et la forme des objets virtuels, de les sélectionner et de les modifier avec un haut degré de réalisme et d'immédiateté. La collision de détection gérée par le système, entre l'instrument virtuel et l'objet, provoque un retour de force sur le bras haptique, ce qui en fait un système plus précis que le gant de données dans le sens où l'instrument peut entrer en contact avec l'objet et ne peut pas le traverser. Dans le cas du gant de données, l'utilisateur peut ressentir les objets mais il peut également les traverser, passer au travers.

Les domaines d'application de ces interfaces concernent par exemple la simulation médicale, les interventions chirurgicales à distance, la modélisation, la manipulation et le dessin en trois dimensions, etc. Bien que les bras mécaniques fussent déjà utilisés dans les années 1940 pour des opérations de manipulation à distance de produits nucléaires, c'est dans les années soixante qu'Alois Knoll présente le premier dispositif haptique permettant de ressentir une forme virtuelle. Par la suite, J.K. Salisbury et Thomas Massie, de la société SensAble Technologies, présentent et commercialisent un bras haptique, le PHANTOM⁶⁷. Ce dispositif inaugura le champ de recherche appelé l'« haptique ».



Figure 3 - Bras haptique « Phantom » de la société SensAble.

⁶⁷ SALISBURY J. Kenneth et SRINIVASAN Mandayam A. *Phantom-Based Haptic Interaction with Virtual Objects*. In *IEEE Computer Graphics Applications*. 17, 1997, pp. 6-10

Limitations des technologies haptiques

L'approche des interactions homme-machine proposée par les technologies haptiques est véritablement en rupture avec les méthodes d'interaction définies par le paradigme WIMP. La notion d'instrument est plus que présente puisque les outils de la manipulation sont physiques et offrent davantage de degrés de liberté que la souris. La non-séparabilité⁶⁸ des degrés de liberté entre l'instrument et la tâche à effectuer augmentent considérablement l'efficacité du dispositif. Cependant, les technologies haptiques souffrent de certaines limitations. D'une part, le contact avec les objets virtuels ne se fait que par un ou des points uniques, ce qui limite la précision de la manipulation de l'objet et la fidélité du retour de force. En effet, le feedback cutané est limité dans la plupart des interfaces haptiques car il s'effectue par l'application d'une force alors que le sens du toucher et la sensation de texture sont normalement ressentis par de multiples micro-déformations de la surface de la peau. Même si aujourd'hui, certains systèmes sont extrêmement perfectionnés et permettent d'interagir avec des objets virtuels de manière très réaliste, ils sont également extrêmement chers. Le coût, l'encombrement et la faible compatibilité de ces systèmes avec les programmes actuels ne favorisent pas une large adoption par les particuliers. Ce sont plutôt les domaines professionnels comme la médecine, les musées ou les designers 3D par exemple qui profitent de la spécialisation de ces outils pour la réalisation de tâches bien spécifiques.

2.4. Instrument mixte

Les interfaces utilisateur tangibles (Tangible User Interface ou TUI) sont des interfaces où les utilisateurs interagissent avec les données numériques à travers l'environnement physique. Hiroshi Ishii, professeur au MIT et directeur du *Tangible Media Group* a été l'un des pionniers avec George Fitzmaurice et William Buxton, en introduisant le concept d'interface palpable (*graspable user interface*) dès 1995. Il s'agit d'interfaces qui permettent de contrôler les objets virtuels grâce à des artefacts physiques appelés *bricks*. Ces briques, de la taille d'une brique *LEGO*, sont posées sur un large écran horizontal et sont manipulables. Un

⁶⁸ JACOB Robert J. K., SIBERT Linda E., McFARLANE Daniel C. et MULLEN M. Preston Jr. *Integrity and separability of input devices*. In *ACM Trans. Computer Human Interaction* (1, 1), 1994, pp. 3-26

objet palpable est défini par Ishii *et al.*⁶⁹ comme un objet constitué d'une partie physique (une ou plusieurs briques) et d'un objet virtuel. Par essence, les interfaces palpables sont un mélange d'artefacts virtuels et physiques où l'affordance des briques physiques facilite la manipulation directe. Les TUI appartiennent au domaine de la réalité mixte car elles mélangent des artefacts physiques et virtuels.

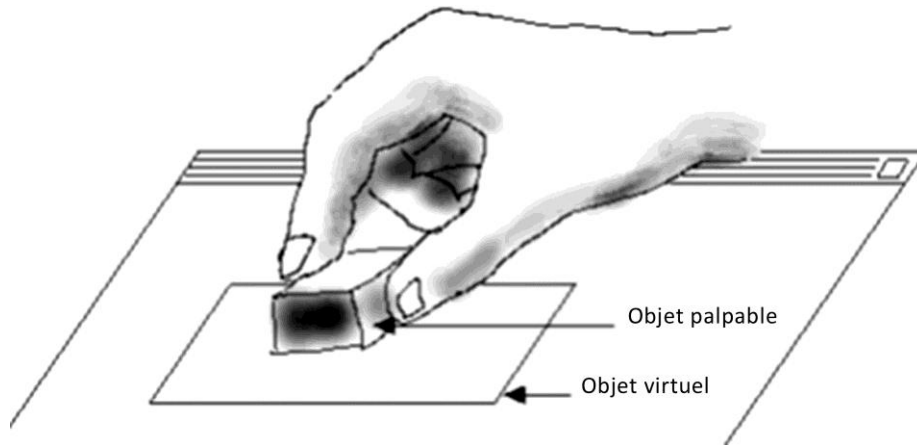


Figure 4 - Représentation du concept d'interface palpable. Reproduction, [Ishii *et al.*, 95] Traduction libre.

Paul Milgram définit les systèmes de réalité mixte comme « *des technologies qui impliquent le mélange de mondes réels et virtuels*⁷⁰ », inscrits sur le continuum de virtualité qui relie les environnements entièrement réels et les environnements entièrement virtuels. Le continuum de virtualité est un ensemble de classes d'interfaces où les environnements réels apparaissent d'un côté du continuum et les environnements virtuels à l'extrême opposée. Tout au long du continuum, nous trouvons la réalité augmentée, situation où l'environnement réel est « augmenté » d'informations ou d'objets virtuels, et la virtualité augmentée, qui fait référence aux interfaces dont l'objet de la tâche réside dans le monde virtuel et où les interactions reposent sur la manipulation d'objets du monde réel.

⁶⁹ FITZMAURICE George W., ISHII Hiroshi et BUXTON William A. S. *Bricks : laying the foundations for graspable user interfaces*. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (CHI '95). ACM Press, New York, USA, pp. 442-449

⁷⁰ MILGRAM Paul et KISHINO Fumio. *A taxonomy of mixed reality visual displays*. In *IEICE Transactions on Information Systems* (Vol. E77-D, No. 12), 1994, 15p.



Figure 5 - Représentation simplifiée du continuum de virtualité.
Reproduction, [Milgram 94]. Traduction libre.

Les TUI s'inscrivent dans le contexte de la virtualité augmentée car dans ce système, l'environnement virtuel est « augmenté » d'une couche de réel. Les instruments de manipulation sont des objets réels et possèdent à la fois des propriétés physiques et numériques.

Les « bits tangibles »

L'expression « interface utilisateur tangible » apparaît en 1997 dans un article⁷¹ écrit par Hiroshi Ishii et Brygg Ullmer. Ils y font état de la vision du *Tangible Media Group* sur les interfaces tangibles. Les auteurs exposent tout d'abord leur intérêt pour les instruments créés par l'homme, bien avant l'invention de l'ordinateur personnel, qui servaient à mesurer le temps, à prédire le mouvement de planètes, à dessiner des formes géométriques ou à calculer. Ils soulignent également s'être inspirés de l'esthétique et de la forte affordance (capacité de l'objet à révéler son utilisation) de ces instruments scientifiques, qui tendent selon eux à disparaître des écoles et des laboratoires, progressivement remplacés par les ordinateurs personnels capables de simuler ces outils sans difficulté. Largement inspiré par l'informatique ubiquitaire, le groupe de recherche du MIT a pour ambition de présenter des moyens concrets pour dépasser le modèle WIMP dominant et le dispositif actuel (écran, clavier, souris). Dans la logique d'une informatique ambiante, ubiquitaire et invisible, le modèle d'interface utilisateur tangible est décrit comme un nouveau type d'interface homme-machine capable d'augmenter le monde réel en couplant des informations numériques aux objets physiques ordinaires, transformant le monde lui-même en interface.

Partant du principe que les hommes ont développé des sens et des compétences sophistiqués pour percevoir l'environnement et pour manipuler des objets physiques, les

⁷¹ ISHII Hiroshi et ULLMER Brygg. *Tangible bits : towards seamless interfaces between people, bits and atoms*. In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '97)*. ACM, New York, USA, 1997, pp. 234-241

interfaces utilisateur tangibles sont conçues afin de profiter de la richesse des capacités de perception multimodales de l'utilisateur et de l'affordance naturelle des objets. La notion de « bits tangibles » décrit l'augmentation de l'environnement virtuel par des objets physiques, saisissables, « palpables », manipulables et contrôlables. L'idée principale est de donner une forme physique aux données numériques (« bits tangibles »). Cette forme physique sert à la fois de représentation et de contrôle pour leur analogon numérique. Ces artefacts physiques sont des instruments appartenant aux deux mondes, réel et virtuel, qui peuvent être spécialisés en fonction de la ou des tâches auxquelles ils sont affectés.

Terminologie

Selon le *Petit Robert*, le terme « tangible » a pour définition : « *Qui tombe sous le sens du tact, que l'on peut connaître en touchant.* » Ce n'est certainement pas suffisant pour caractériser les interfaces utilisateur tangibles. Car au-delà de faire référence au caractère « palpable » de l'interface, par l'existence d'artefacts, les interacteurs tangibles, dans le monde physique des objets, le terme « tangible » renvoie également au caractère de « *quelque chose dont la réalité est évidente* », comme si l'interpénétration du réel et du virtuel, par l'intermédiaire de ces artefacts tactiles et préhensibles, constituait un nouvel espace d'actualisation dont la nature s'inscrivait dans une réalité qui s'adresse autant au corps qu'à l'esprit. La distinction entre la réalité physique et la réalité virtuelle s'estompe pour ne former qu'une seule et même réalité, tangible, générée par des artefacts physiques qui se trouvent augmentés de nouvelles dimensions (positions, temporalité et interactivité virtuelles). Manipuler ces artefacts constitue l'expérience nécessaire à l'acceptation d'une réalité dont le modèle suggère de nouveaux archétypes, de nouveaux espaces d'action et d'expérience. Le feedback tactile de ces objets rappelle à l'esprit qu'ils constituent une réalité : la main se sent touchée aussi bien qu'elle touche. Les interfaces tangibles, contrairement aux interfaces tactiles, tiennent davantage de la réalité que de l'illusion, en ce sens où elles s'imposent en continue dans l'espace de notre perception sensible. C'est l'idée d'une réalité tangible proche de celle de l'écrivain Philip K. Dick pour qui « *la réalité, c'est ce qui continue à s'imposer à vous quand vous cessez d'y croire.*⁷² »

En plus de fournir des informations sur l'objet virtuel qu'ils représentent, les interacteurs tangibles offrent à l'utilisateur des capacités d'action sur le monde réel et virtuel. Ils sont à la

⁷² DICK Philip K. *La trilogie divine, Tome 1 : Siva*. Editions Gallimard 2006, 353 p. (Coll. Folio SF)

fois objet de représentation et outil de contrôle des données numériques. De la même manière qu'un objet réel est composé d'une série de propriétés, un objet numérique relève de paramètres et de données numériques. Un objet mixte est composé des deux : des données physiques liées aux données numériques. Céline Coutrix et Laurence Nigay ont proposé en 2006 un modèle d'interaction mixte⁷³ où elles considèrent que pour décrire le lien entre les propriétés physiques et numériques d'un objet mixte, il fallait considérer deux niveaux de modalité. Les modalités qui définissent le lien entre les propriétés physiques et numériques d'un interacteur tangible sont appelées les modalités de couplage (*linking modalities*) en opposition aux modalités d'interaction qui sont utilisées par l'utilisateur pour interagir avec l'environnement mixte.

Les TUI se différencient des GUI en ce sens où l'utilisateur interagit avec le système, non plus de manière indirecte mais en manipulant les représentations tangibles des données numériques. L'utilisateur peut ainsi communiquer avec le système en manipulant de ses deux mains les formes tangibles et intangibles de la représentation numérique. D'un côté, les représentations tangibles permettent de créer des liens entre l'environnement réel et l'environnement virtuel selon des modalités de couplage et d'interaction. Ce sont des objets physiques qui s'incarnent soit sous la forme d'instruments (ex : lentilles passives, *MetaDesk*. cf. Figure 6), on les appellera alors des instruments mixtes afin de prolonger le modèle d'interaction instrumentale vu précédemment, ou dans la métaphore d'un objet d'intérêt (ex : synthétiseurs modulaire, *ReacTable*, cf. Figure 9). Dans les deux cas, ils sont utilisables comme des objets réels car ils sont physiquement préhensibles. D'un autre côté, les représentations intangibles permettent de montrer des objets ou des notions abstraites qui n'existent pas dans la réalité. En effet, les représentations intangibles correspondent aux visualisations générées sur l'écran et peuvent donc prendre des formes impossibles et subir toutes sortes de transformations que la physique des objets ne permet pas.

⁷³ COUTRIX Céline et NIGAY Laurence. *Mixed reality: a model of mixed interaction*. In *Proceedings of the working conference on Advanced visual interfaces* (AVI '06). ACM, New York, USA, pp. 43-50

Caractérisation des interfaces tangibles

Hiroshi Ishii⁷⁴ définit trois caractéristiques majeures pour définir une interface utilisateur tangible :

- *Le couplage informatique des représentations tangibles aux données numériques sous-jacentes.* C'est la caractéristique principale des TUI de définir des modalités de couplage entre les représentations physiques et numériques des objets mixtes. Les objets peuvent être de formes et de tailles variables et plus ou moins abstraites du moment que le couplage réel/virtuel fournit suffisamment de données numériques et de feedback pour compléter les données provenant du réel.
- *Les mécanismes de contrôle interactif des représentations tangibles.* Les interacteurs tangibles doivent être inertes et ne bouger qu'à condition que l'utilisateur les manipule directement. Il est également possible que ces objets tangibles soient actualisés physiquement ou à retour de force. La manipulation de ces artefacts doit pouvoir se faire librement, dans les six degrés de liberté permis par l'espace même s'il est possible de les contraindre faiblement sur une surface lors de leur déplacement par exemple. Enfin, l'interaction attendue par le système doit être réalisée grâce à l'interacteur tangible avec un mécanisme simple et bien connu de l'utilisateur pour simplifier les interactions et les rendre plus faciles à apprendre.
- *Le couplage perceptuel des représentations tangibles aux représentations intangibles dynamiques.* Les interfaces utilisateur tangibles fonctionnent sur l'équilibre entre les objets tangibles et intangibles. Bien que les interacteurs tangibles jouent un rôle central dans la représentation et le contrôle des TUI, les représentations intangibles, éléments graphiques ou sonores, participent grandement au dynamisme du couplage entre les deux. D'une part, les représentations intangibles permettent de créer une continuité spatiale entre l'espace tangible et intangible de l'interface, assurant le couplage perceptuel des représentations. D'autre part, le feedback en temps réel provoqué par la manipulation d'un interacteur tangible garantit l'effet de causalité d'un espace à l'autre.

⁷⁴ ISHII Hiroshi. *Tangible bits : beyond pixels*. In *Proceedings of the 2nd international conference on Tangible and embedded interaction* (TEI '08). ACM, New York, USA, 2008, chap. XV-XXV

Exemple 1 : MetaDesk, 1997

MetaDESK est une surface interactive tangible développée en 1997 par Hiroshi Ishii et Brygg Ullmer du MIT afin de mettre en pratique et d'explorer les interfaces utilisateur tangibles. Il s'agit d'un dispositif composé d'un bureau, d'un écran à rétroprojection quasi horizontal, d'un écran plat monté sur un bras mécanique appelé « lentille active » et d'une « lentille passive » transparente qui est posée sur l'écran de projection.

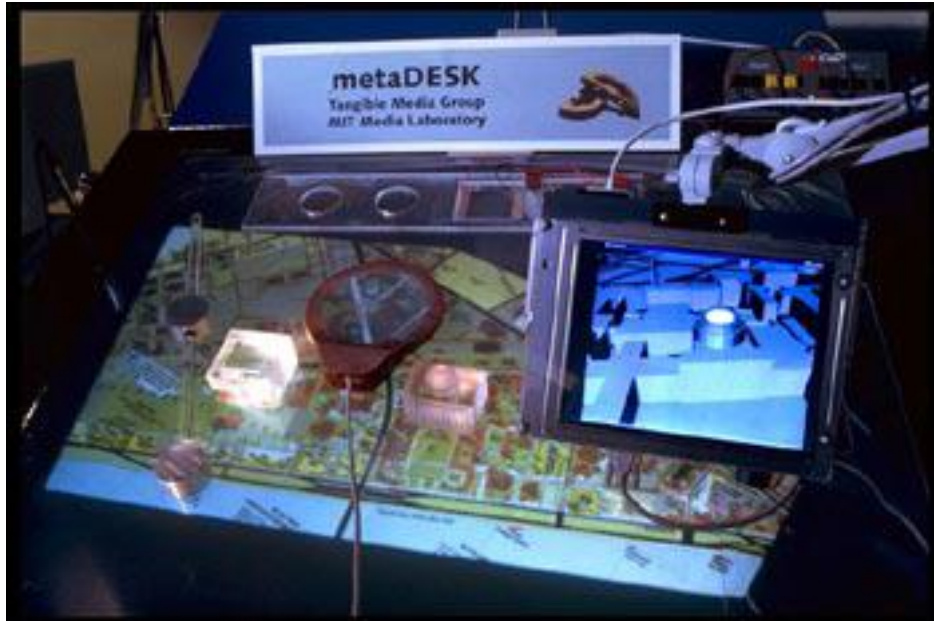


Figure 6 - Dispositif metaDESK. In [Ishii et Ullmer, 1997]

La première application présentée sur le *metaDESK* est *Tangible Geospace*⁷⁵, en 1997, un logiciel qui permet à l'utilisateur de naviguer à travers les cartes géographiques et les univers en 3D représentant les environs du campus du MIT. En déplaçant la « lentille active », l'utilisateur peut naviguer dans l'espace en trois dimensions. La « lentille passive » est déplaçable sur la surface de l'écran et agit comme les « lentilles magiques » utilisées dans les *ToolGlass* [Bier *et al.* 1993; Stone *et al.* 1994] en affichant par exemple une vue satellite de la carte sous la lentille ou en montrant le bâti passé ou futur.

Afin de manipuler la carte géographique, *metaDESK* propose d'utiliser des interacteurs tangibles qui sont l'incarnation des métaphores du bureau utilisées au sein des interfaces WIMP. Les *phicones* sont par exemple des incarnations physiques des icônes qui, une fois

⁷⁵ ISHII Hiroshi et ULLMER Brygg. *The metaDESK : models and prototypes for tangible user interfaces*. In *Proceedings of the 10th annual ACM symposium on User interface software and technology (UIST '97)*. ACM, New York, USA, pp. 223-232

positionnées sur la surface de l'écran, sont repérées par les capteurs électromagnétiques. Dans l'application *Tangible Geospace*, les *phicones* représentent des bâtiments et lorsqu'elles sont posées sur l'écran, la carte vient s'ajuster automatiquement en fonction de leur position.

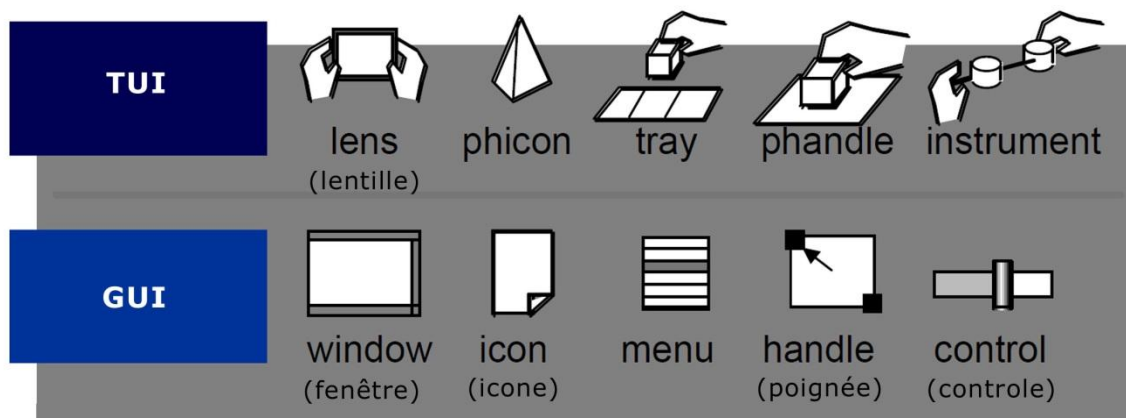


Figure 7 - Incarnation physique des éléments du GUI dans les TUIs. In [Ishii et Ullmer, 97]. Traduction libre.

Exemple 2 : ReactTable - 2005

Le projet *Reactable* est né de la volonté de ses concepteurs Sergi Jordà, Günter Geiger et Martin Kaltenbrunner, de réaliser un instrument de musique électronique d'un nouveau genre qui transforme la pratique de l'instrument de musique en une expérience tangible et visuelle. Selon Jordà *et al.*⁷⁶, l'idée principale est de créer « *un instrument de musique électronique tangible qui permet à des musiciens de réaliser des performances collaboratives live sans les limites de nombreuses interfaces dédiées à la musique électronique.* » La *Reactable* est une surface de projection horizontale et circulaire, une table interactive qui permet à une ou plusieurs personnes de jouer en même temps par l'intermédiaire d'un jeu d'interacteurs disposés sur la surface. En mettant ces interacteurs en situation de proximité les uns avec les autres, l'utilisateur construit et joue d'un instrument en même temps. L'implication simultanée des sens de la perception, la forte affordance des objets mixtes ainsi que le feedback visuel généré par la projection vidéo créent les conditions d'une expérience de création/interprétation musicale multimodale et synesthésique. Présentée

⁷⁶ « *The main idea is the creation of a tangible electronic musical instrument that allows expressive collaborative live performances for professional musicians without the limits of many screen-based interfaces for electronic music* » in KALTENBRUNNER Martin, GEIGER Günter et JORDÀ Sergi. *Dynamic patches for live musical performance*. In *Proceedings of the 2004 conference on new interfaces for musical expression* (NIME '04), Michael J. Lyons. National University of Singapore, Singapore, 2004, pp. 19-22. Traduction libre : COQUELIN Marine et MORIN Rudy.

lors de festivals de création numériques ou lors de conférences comme *Ars Electronica*, le *Sónar* ou encore le *SIGGRAPH*, la *Reactable* est un exemple d'application des TUI récent qui démontre bien la rupture de paradigme qu'elles produisent et les possibilités nouvelles d'interaction qu'elles mettent à disposition du designer d'interface.

La zone d'interaction de la *Reactable* est une surface semi-transparente sur laquelle sont projetées les représentations des propriétés numériques et des modalités de couplage des interacteurs tangibles posés sur l'interface. Une caméra vidéo située dessous, analyse par l'intermédiaire d'une technologie optique infrarouge (illuminateurs et caméra IR) l'activité de la surface de la table, repérant la nature, la position et l'orientation des interacteurs. Une fois posés sur la surface, ces objets s'illuminent et la génération de la musique électronique s'affiche sur la surface, représentée par des lignes partant des interacteurs générateurs et allant jusqu'aux interacteurs de sorties, par des ondulations et des cercles de lumière qui correspondent aux différents types d'ondes ou traitements sonores appliqués au signal.



Figure 8 - Reactable : instrument tangible interactif. Jordà *et al.*, 2005.

Ces interacteurs tangibles sont identifiables par leur forme et par les pictogrammes imprimés sur leur surface qui indiquent la nature de leur fonction. Ce sont les représentations physiques des composants traditionnels des synthétiseurs modulaires constitués de différents générateurs sonores. Tandis que certaines formes comme les carrés

ou les cubes représentent des générateurs et des filtres sonores, d'autres comme les pentagones représentent des filtres et des modulateurs. En déplaçant ces instruments mixtes, en modifiant leur distance, leur orientation et leurs relations, ces actions modifient la structure topologique et les paramètres de la synthétisation du son. Lorsque ces interacteurs sont disposés sur la surface, ils sont entourés de cercles blancs permettant d'ajuster, grâce au contact du doigt sur la surface, l'amplitude de l'effet attendu par l'interacteur, comme le volume ou l'intensité de l'application d'un filtre sonore.

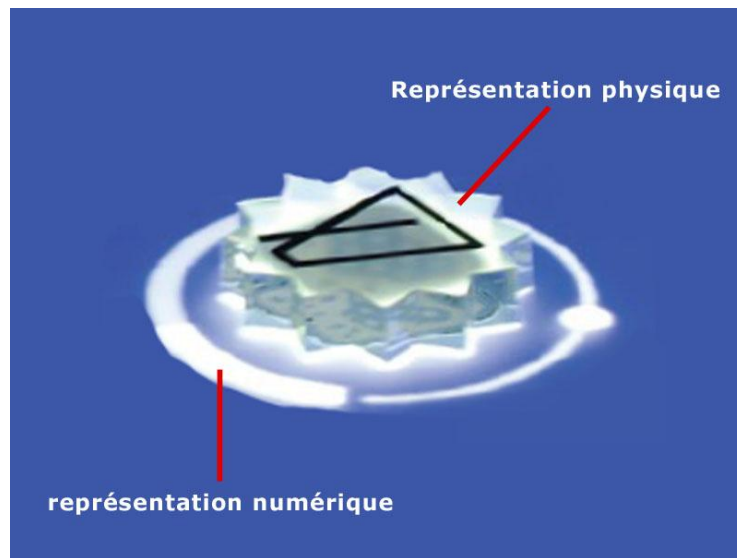


Figure 9 - Interacteur tangible mixte utilisé sur la Reactable.

La projection sur la surface dessine le cheminement des sons synthétisés, les propriétés numériques des interacteurs et les interactions entre les différents filtres. L'affichage réactif de ces informations numériques et la représentation graphique des activités du système fournissent un feedback visuel des états et des principales caractéristiques de la génération et du flux sonore, feedback indispensable à la jouabilité de l'instrument. Le système est également gouverné par les règles de couplage entre les propriétés physiques et numériques des interacteurs mais également par des règles de connexion de ces objets mixtes. En effet, le principe de connexion dynamique sur la *Reactable* n'oblige pas l'utilisateur à connecter explicitement les objets tangibles entre eux. Il s'agit de règles simples qui connectent ou déconnectent les objets en fonction de la distance qui les sépare aux autres. Déplacer l'interacteur sur la surface de la table produit obligatoirement l'altération et la modulation des connexions existantes. Afin d'éviter toute interférence avec la topologie existante, deux moyens existent. Le premier est de retirer l'interacteur tangible

de la surface, en le soulevant, ce qui le désactive et supprime les effets provoqués par cet interacteur sur les interacteurs à proximité. Le second consiste pour l'utilisateur à créer des connexions permanentes en reliant deux interacteurs par un geste de liaison. L'effet persiste tant que l'un des deux interacteurs n'est pas retiré de la surface.



Figure 10 - La Reactable en cours d'utilisation.

La *Reactable* réunit toutes les caractéristiques définies par Hiroshi Ishii concernant les TUI. Le couplage des représentations physiques aux données numériques sous-jacentes se trouve ici représenté par la métaphore du flux sonore. Les interacteurs physiques ont des formes peu spécialisées, contrairement au projet *metaDESK* où ils étaient l'incarnation des instruments virtuels communs aux interfaces WIMP. Toutefois, les interacteurs se distinguent d'une part grâce à leur forme (figure géométrique simple) qui indique le type de module du synthétiseur et d'autre part par le pictogramme dessiné sur la face visible qui renseigne sur la fonction propre à ce module. Les mécanismes de contrôle interactif des représentations tangibles est facilité par le feedback visuel et sonore de l'interface mais également par le système de connexion dynamique qui définit les relations entre plusieurs interacteurs grâce à une règle simplement basée sur la distance et qui est facilement comprise par l'utilisateur. Enfin, le couplage perceptuel des représentations est assuré par l'affichage des propriétés numériques des interacteurs tout autour de l'objet physique permettant de créer une continuité spatiale entre l'espace tangible et intangible.

Limitation des TUI

Une des caractéristiques importantes des interacteurs mixtes tient dans leur coexistence dans les deux espaces de l'interaction tangible, l'espace réel et l'espace virtuel. Ces artefacts possèdent des propriétés physiques mais également des propriétés numériques dont l'actualisation s'effectue selon les modalités de couplage définies par le designer d'interface. La forte affordance de ces objets physiques permet à l'utilisateur de les prendre en main et de les manipuler de manière naturelle ou d'une manière que ces objets inspirent d'emblée. Les TUI envisagent les interactions homme-machine en incarnant les instruments de la manipulation sous la forme d'objets mixtes dont le couplage avec l'interface et la superposition physique de l'instrument sur la surface d'affichage permet de respecter davantage les principes de la manipulation directe que les interfaces classiques.

Cependant, en considérant les interacteurs tangibles comme des artefacts physiques, des objets persistants que nous pouvons certes poser sur la surface d'affichage et soulever mais qui ne peuvent être créés dynamiquement ou disparaître spontanément comme les fenêtres graphiques ou les icônes, nous serions en droit de soulever des problématiques d'ordre pratique comme l'encombrement des interacteurs tangibles. En effet, la multiplication des instruments d'interaction peut devenir préjudiciable à la visibilité de l'interface et de l'objet d'intérêt sur lequel s'applique la manipulation. Fitzmaurice⁷⁷ recommande justement l'utilisation de multiples interacteurs tangibles et notamment de préférer utiliser des objets physiques spécialisés pour interagir avec le système, car d'une part les interacteurs spécialisés sont taillés pour des tâches spécifiques et d'autre part, une collection d'objets encourage l'interaction bi-manuelle et multiutilisateur.

De plus, ce type d'interface dispose d'entrées multiplexées dans l'espace et non dans le temps, ce qui augmente les possibilités d'interférences entre les interacteurs. Dans le cas de metaDESK par exemple, Brygg Ullmer and Hiroshi Ishii⁷⁸ évoquent des ambiguïtés d'interaction lorsque deux interacteurs géo-référents, le *dôme* et la *phicone* du *Media Lab*, sont manipulés en même temps. Leur solution est de créer un nouvel instrument, composé de deux cylindres couplés mécaniquement par une réglette afin de contraindre la rotation

⁷⁷ FITZMAURICE George W. *Graspable User Interfaces*. Thèse, University of Toronto, Toronto, Canada, Canada. Sous la direction de William Buxton, 1996

⁷⁸ ISHII Hiroshi et ULLMER Brygg. *The metaDESK : models and prototypes for tangible user interfaces*. In *Proceedings of the 10th annual ACM symposium on User interface software and technology (UIST '97)*. ACM, New York, USA, pp. 223-232

des deux interacteurs et de prévenir les « interactions ambiguës » par une contrainte mécanique.

Enfin, les TUI sont elles-mêmes spécialisées et leur application ne concerne que certains domaines industriels, artistiques ou culturels et peinent à rentrer dans les usages courants de l'informatique, et plus encore auprès des particuliers. D'une part, les TUI sont souvent d'une taille suffisamment grande pour y poser les interacteurs et disposer d'une surface interactive bénéficiant d'interactions bi-manuelles et multiutilisateurs. La physique des objets posés sur la surface implique généralement une surface horizontale ou légèrement inclinée qui contraint davantage les usages autorisés par ces technologies.

CHAPITRE II INTERFACES TACTILES MULTI-POINTS

« Les images savent désormais qu'on les regarde. »

Richard Bolt

Introduction

Interfaces Utilisateur Naturelles (NUI⁷⁹)

Au cours des chapitres précédents, nous avons pu observer la richesse et la diversité des recherches en IHM menées depuis ces cinquante dernières années en mettant en regard l'évolution de la notion d'instrument et les modalités de couplage entre le périphérique et l'interface. D'un côté du continuum de virtualité, les instruments haptiques s'incarnent intégralement dans le réel de manière prothétique et procurent à l'utilisateur la conviction de la matérialité des objets virtuels par le réalisme des retours tactilo-kinesthésiques. Plus loin, les interacteurs mixtes, artefacts physiques de couplage entre l'espace réel et virtuel, profitent de l'affordance de leur forme physique et de leur spécialisation pour garantir le réalisme des interactions. De l'autre côté du continuum, les interfaces tactiles déplacent la partie physique de l'instrument d'interaction vers la main de l'utilisateur, supprimant ainsi toute possibilité de préhension des instruments d'interaction ou des objets de la manipulation. Les interfaces utilisateur tactiles multi-points, qui intéressent notre étude, et dont la manipulation est essentiellement manuelle, présente un aspect singulier de l'instrument d'interaction. Nous verrons dans quelle mesure la main de l'utilisateur, grâce à la richesse fonctionnelle du geste, acquiert sa dimension instrumentale.

En juillet 2006, Christian Moore fonde un groupe d'intérêt⁸⁰, le *NUI Group*, qui réunit plusieurs milliers de membres (étudiants, chercheurs, artistes numériques et amateurs) et dont la motivation est de développer des NUI sous licence libre, d'expérimenter et de partager les résultats de leur recherche au plus grand nombre. Il propose de nommer « interface utilisateur naturelle » (*Natural User Interface* ou NUI) les interfaces caractérisées par l'« invisibilité », la transparence de leurs instruments d'interaction, par la facilité d'utilisation de l'interface par un novice et la relative simplicité ou transparence de l'apprentissage. Cette définition s'inscrit dans les fondements de l'informatique ubiquitaire tout en intégrant la notion d'interactions dites « naturelles ». Dû à l'ambiguïté soulevée par ce terme, nous verrons au cours des chapitres suivants ce qui relève de cette propriété et

⁷⁹ *Natural User Interface*

⁸⁰ NUI Group. Disponible ici : <http://www.nuigroup.com/>

dans quelle mesure elle est impliquée dans les interactions tactiles multi-points. Tout comme les interfaces multimodales, les interfaces utilisateurs tactiles s'inscrivent dans cette branche de la recherche en IHM. Au cours de ces dernières années, elles ont gagné la confiance des industriels, suite à l'affinage de la technologie, et des particuliers, séduits par la promesse d'une utilisation simplifiée. Nous nous attarderons sur l'histoire de cette technologie, en évoquant les étapes successives de son invention, de son affinage technique et de son adoption, correspondant à la longue période d'innovation qui a précédé sa popularité toute récente.

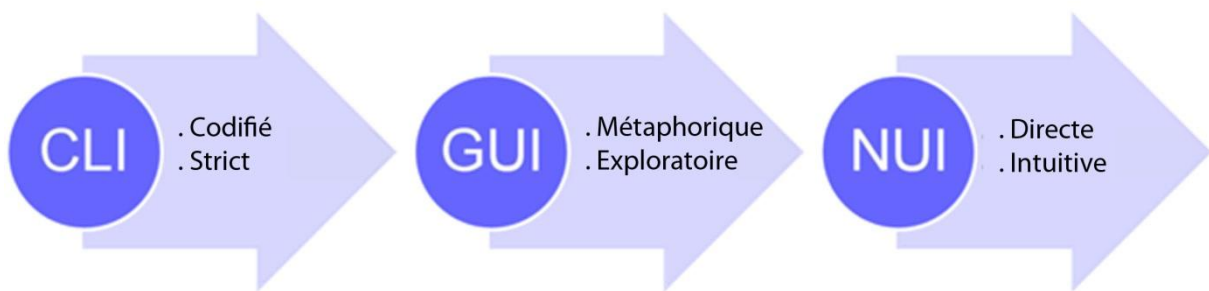


Figure 11 - Evolution simplifiée des IHM. De gauche à droite, Interface en Ligne de Commande (CLI), Interface graphique Utilisateur (GUI), Interface utilisateur Naturelle (NUI).

Informatique ubiquitaire

L'informatique ubiquitaire est un concept, ou plutôt une vision, où le traitement d'information est intégré de manière transparente aux objets et aux activités de notre quotidien. L'idée principale suggère que l'utilisation de l'informatique ubiquitaire implique de nombreux dispositifs et de systèmes simultanément sans que l'utilisateur en soit réellement conscient. Dans ce modèle, l'ordinateur devient une force utile mais invisible qui aide l'utilisateur à réaliser des tâches de son quotidien sans se mettre au travers de son chemin. Ce paradigme est également appelé « intelligence ambiante » ou « informatique omniprésente » (*Ubiquitous Computing*). L'informatique ubiquitaire peut être définie comme l'opposé de la réalité virtuelle. Là où la réalité virtuelle projette les utilisateurs à l'intérieur d'un monde numérique généré par l'ordinateur, l'informatique ubiquitaire force les ordinateurs à s'intégrer dans notre monde de manière invisible. Invisible en ce sens où les systèmes informatiques sont tellement bien intégrés, adaptés et devenus ordinaires, que nous pouvons les utiliser sans même y penser. Ce concept marque une véritable rupture

avec les modèles traditionnels ou haptiques qui exposent l'interface physique pour donner l'intuition de son utilisation à l'utilisateur.

Mark Weiser propose le terme d'informatique ubiquitaire en 1988 lors de ses fonctions en tant que chef de la recherche à Xerox. En reconnaissant que le développement du traitement de l'information dans les scénarios de tous les jours implique une compréhension plus large des phénomènes sociaux, culturels et psychologiques, il s'inspire de la philosophie, de la phénoménologie, de l'anthropologie, de la psychologie et de la sociologie en proposant un modèle d'interaction post-WIMP, centré utilisateur, et qui marque la voie de la pénétration des « *choses qui pensent* » (*things that think*) dans notre quotidien.

Dans un article pour le magazine *Scientific American*, Mark Weiser expose son concept :

« L'informatique ubiquitaire fait référence à la troisième vague informatique, qui n'en est qu'à ses débuts. Au départ il y avait les ordinateurs centraux (ou mainframes), chacun partagé par beaucoup de monde. Aujourd'hui nous sommes dans l'air des ordinateurs personnels où les gens et les machines se confrontent avec un certain malaise par l'intermédiaire d'un bureau. Ensuite vient l'informatique ubiquitaire, ou l'ère des technologies calmes, où la technologie s'estompe à l'arrière-plan de notre vie.⁸¹ »

Il poursuit son article en proposant trois formes de systèmes ubiquitaires, de dimensions différentes, qu'il nomme *tabs*, *pads* et *boards*. En s'attardant sur les descriptions données par Weiser, il nous sera possible de reconnaître l'extrême justesse de sa vision lorsque nous comparerons ces trois dispositifs aux périphériques actuels : les téléphones mobiles, les tablettes et les tables interactives. Nous verrons combien la miniaturisation des composants électroniques, la portabilité des dispositifs, la disparition de périphériques d'entrée physiques et le degré d'abstraction des interactions impliquées dans les surfaces multi-tactiles reflètent de manière prégnante les idées avancées par Mark Weiser vingt ans plus tôt.

⁸¹ « *Ubiquitous computing names the third wave in computing, just now beginning. First were mainframes, each shared by lots of people. Now we are in the personal computing era, person and machine staring uneasily at each other across the desktop. Next comes ubiquitous computing, or the age of calm technology, when technology recedes into the background of our lives.* » In WEISER Mark. *The computer for the 21st century*. In *Human-computer interaction*, Ronald M. Baecker, Jonathan Grudin, William A. S. Buxton, and Saul Greenberg . San Francisco, USA, Weiser, pp. 933-940. Traduction libre : COQUELIN Marine et MORIN Rudy.

1. Une longue période d'incubation

Les technologies multi-tactiles ont une longue histoire et sont restées jusqu'à récemment cloisonnées aux sphères de la recherche ou aux milieux spécifiques comme le domaine artistique ou musical. Ce n'est que depuis quelques années, avec notamment l'introduction de l'*iPhone* d'*Apple* sur le marché, qu'un véritable engouement de la part des industriels et des particuliers est né autour des possibilités offertes par cette technologie. Cependant, *Apple* n'a fait que porter une technologie existante dans un produit largement diffusé comme le téléphone portable, en améliorant le dispositif et en reprenant des principes d'interaction déjà présents dans la littérature scientifique. Par exemple, le geste de pincement (*pinch*) permettant de zoomer sur une image a été présenté initialement par Myron Krueger de la société *Artificial Reality* dès 1982 dans le projet *Videodesk*, un système consistant en une large surface interactive sur laquelle on utilise le mouvement des bras, des mains et des doigts afin d'interagir avec le système. La question n'est pas tant de remettre en cause la propriété intellectuelle de la technologie multi-tactile ou de savoir qui a inventé les modalités gestuelles existantes mais plutôt de mettre en perspective l'évolution et la lente démocratisation de cette technologie au cours de ces trente dernières années.

L'invention de la souris et le temps d'absorption nécessaire aux industriels est une illustration assez révélatrice du phénomène de diffusion d'un concept et des processus impliqués dans l'innovation. En effet, la souris, initialement créée vers 1965 par William English et Doug Engelbart, a été utilisée dans un premier temps par d'autres chercheurs puis adoptée par le laboratoire PARC de *Xerox* vers 1973. Ce n'est qu'en 1984, avec le premier *Macintosh*, que la souris a été proposée au grand public. Puis il a fallu attendre 1995, avec l'arrivée de *Microsoft Windows 95*, pour qu'elle devienne omniprésente et qu'elle soit adoptée par un large public. Les avantages d'un tel dispositif d'interaction semblent évidents et nous serions en mesure de nous demander pourquoi il a fallu attendre trente ans entre son invention et son utilisation systématique. Le temps d'incubation nécessaire à un type d'interface homme-machine avant sa pénétration et sa diffusion auprès du grand public peut s'expliquer par de nombreux facteurs comme les coûts de fabrication, l'amélioration du dispositif, les temps d'apprentissage et l'adoption progressive de la technologie par les utilisateurs. Dans un article de 2008, Bill Buxton reprend et inverse le concept de « longue traîne » (*long tail*), présenté en 2004 par Chris Anderson qui désigne un phénomène connu

par les statisticiens décrivant les lois de distribution d'un produit en fonction des canaux de distribution qui le rendent accessible. Il propose le concept de « *long nose of innovation* » (littéralement : le long nez de l'innovation) qui décrit la longue et discrète période de maturation d'une technologie avant sa démocratisation et son adhésion par le plus grand nombre.

« Je suis convaincu qu'il existe un concept miroir de la « longue traîne » qui est aussi important pour ceux qui souhaitent comprendre les processus d'innovation. Il explique que la majeure partie de l'innovation avant le moment d'exclamation (le multi-tactile sur iPhone par exemple) est également de faible amplitude et s'étale sur une longue période - Avant que la « nouvelle » idée soit devenue largement connue et bien avant qu'elle atteigne le point de bascule. C'est ce que j'appelle le Long Nez de l'Innovation.⁸² »

Selon Bill Buxton, le processus d'innovation procède d'une longue maturation, de l'idée initiale jusqu'à son adoption par les utilisateurs. L'invention passe par une période d'analyse, de mise en situation, d'amélioration et d'affinage qui constitue le temps de recherche et de développement lié à l'émergence d'une nouvelle technologie.

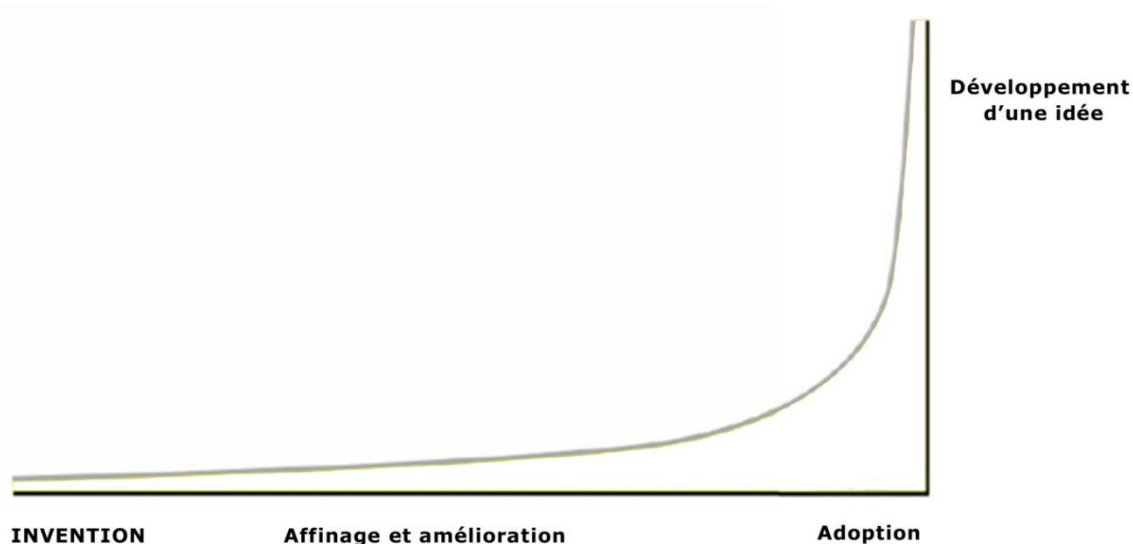


Figure 12 - Les phases du processus d'innovation en fonction du temps. Reproduction du graphique. In [Buxton Bill, 2008].

⁸² « My belief is there is a mirror-image of the long tail that is equally important to those wanting to understand the process of innovation. It states that the bulk of innovation behind the latest “wow” moment (multi-touch on the iPhone, for example) is also low-amplitude and takes place over a long period—but well before the “new” idea has become generally known, much less reached the tipping point. It is what I call The Long Nose of Innovation. » In BUXTON Bill. *The Long Nose of Innovation*. In *Businessweek*, Janvier 2008

2. Histoire des technologies multi-tactiles

Les travaux et les publications sur les technologies tactiles, multi-tactiles et gestuelles ont une longue histoire et il me semble pertinent de mettre en perspective l'évolution de ces technologies afin de comprendre leurs origines et les concepts qu'elles véhiculent depuis une quarantaine d'années. Bien qu'aujourd'hui, la prolifération d'appareils ou de dispositifs pourvus d'une technologie tactile multipoints semblent suggérer l'adoption d'une technologie récente, les premières explorations et développements scientifiques remontent aux années soixante.

1960 – 1991 : une phase expérimentale

C'est justement au cours des années soixante qu'apparaissent les premières interfaces dont l'écran tactile permet d'interagir avec le système. Il s'agit de travaux de recherche⁸³ menés par IBM en collaboration avec l'université d'Illinois. Ces ordinateurs quittent les laboratoires d'IBM pour être installés dans des universités en **1972** sous l'apparence d'ordinateurs personnels et sont déjà connectés à un réseau local. Connus sous le nom de *PLATO IV*, ces terminaux sont équipés d'un écran plasma et d'une surface tactile. À cette même période, le laboratoire PARC de Xerox commence seulement à travailler sur le projet *ALTO*.



Figure 13 - Terminal PLATO IV, 1972.

⁸³ National Research Council. *Innovation in Information Technology*. In *Computer Science and Telecommunications Board*. Washington DC, The National Academies Press, 2003, 84 p.

Les années **1970** marquent le développement de technologies tactiles comme les écrans capacitifs, résistifs ou à onde de surface, qui sont par ailleurs toujours utilisés aujourd'hui. Cependant, les recherches et les développements de ces technologies sont menés et commercialisés par des sociétés comme *Elographics* ou *MicroTouch System* et interviennent à un moment où le design d'interface graphique tend à privilégier la souris comme mode d'interaction. Ces écrans tactiles manquent de précision et ont des temps de réponse relativement longs. Durant une longue période, ces écrans tactiles à un seul point de contact, auront seulement servi à remplacer la souris comme dispositif de pointage et n'auront été utilisés que marginalement dans des bornes d'affichage, des distributeurs de billets (*SNCF* par exemple) ou des dispositifs spécifiques où la souris est peu pratique (dans un musée par exemple).

Les premières apparitions du concept de technologie multi-tactile pour le contrôle manuel de systèmes numériques figurent en **1982** dans la thèse⁸⁴ de Nimish Mehta, étudiant à l'université de Toronto. L'interface machine flexible (*Flexible Machine Interface*) présentée par ce doctorant est un système à reconnaissance optique qui utilise une caméra pour capturer les ombres des doigts à travers une surface translucide. Capable de reconnaître des formes et la pression des doigts exercée sur la surface, le système mis au point par Nimish Mehta est précurseur d'un bon nombre de systèmes utilisés encore aujourd'hui comme la table interactive *Microsoft Surface V1* qui utilise des capteurs optiques infrarouges pour repérer le contact des doigts à travers une surface de projection en polycarbonate (*Altuglass*) translucide.

Dès **1983**, Myron Kruger fait la démonstration du système *Videodesk*, une surface interactive qui réagit aux mouvements des bras, des mains et des doigts. Le dispositif est pourvu d'une caméra installée au-dessus de la surface qui capture les mouvements de l'utilisateur et les transmet à un ordinateur. Le système analyse les positions et les actions capturées. Les formes créées par les bras et les mains de l'utilisateur sont analysées et représentées à l'écran sous la forme de contours. Dans ce dispositif, l'écran vertical est situé devant l'utilisateur alors que la surface de manipulation est horizontale et posée sur le bureau. Cette séparation s'explique par les moyens limités de l'époque mais les concepteurs imaginent déjà des systèmes plus naturels où la surface d'affichage est en même temps la

⁸⁴ MEHTA Nimish. *A Flexible Machine Interface*. Thèse, Department of Electrical Engineering, University of Toronto, 1982

surface d'entrée gestuelle. Afin de simplifier les interactions utilisateur, les silhouettes du bras et de la main sont représentées à l'écran pour donner des indications de position précises des manipulations.

Parmi les applications présentées sur ce dispositif, certaines montrent des interactions simples et intuitives comme dessiner à l'écran avec le doigt tandis que d'autres applications plus innovantes font intervenir des gestes de commande spécifiques comme effacer un dessin en écartant tous les doigts de la main. Ce système est précurseur dans le sens où l'approche gestuelle présente les premières formes d'interactions réalisées par des gestes simples et intuitifs comme le pointage, le déplacement et le pincement d'objets (*pinch*) pour en modifier respectivement la position et la taille.

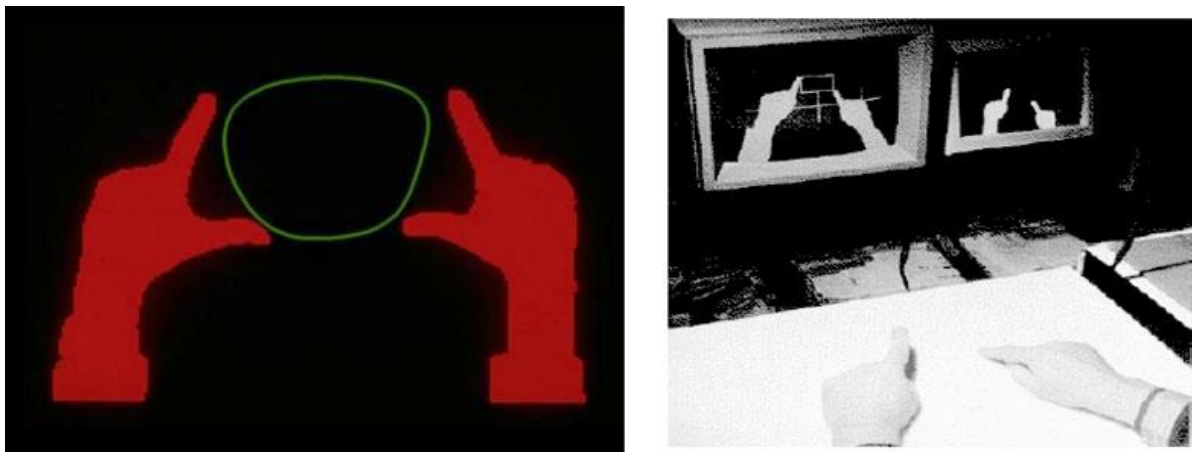


Figure 14 - Videodesk de Myron Krueger, 1983. À gauche un exemple de geste de pincement. À droite une vue du dispositif.

En **1984**, les laboratoires *Bell Labs* développent un écran tactile qui permet de manipuler une image avec deux mains. La même année, Bill Buxton et son équipe⁸⁵ développent une tablette tactile capable de sentir différents degrés de pression de manière continue et cela pour chacun des points de contact. Leurs travaux s'inscrivent directement dans la lignée des recherches menées précédemment par Sasakl *et al.*⁸⁶ en 1981 et Nimish Metha⁸⁷ en 1982. En **1985**, après avoir présenté leur tablette au SIGCHI, groupe d'intérêt sur l'interaction homme-machine, Bill Buxton est contacté par Lloyd Nakatani qui l'invite à visiter les

⁸⁵ LEE SK, BUXTON William et SMITH K. C. *A multi-touch three dimensional touch-sensitive tablet*. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (CHI '85). ACM, New York, USA, 1985, pp. 21-25

⁸⁶ SASAKL L., Fedorkow, G., BUXTON William, RETTARETH C., et Smith K.C. *A Touch-Sensitive Input Device*. In *Proceedings of the Fifth International Conference on Computer Music*. North Texas State University, Texas, 1981

⁸⁷ MEHTA Nimish. *A Flexible Machine Interface*. Thèse, Department of Electrical Engineering, University of Toronto, 1982

laboratoires *Bell* et découvrir la technologie développée par Bob Boie. Il s'agit d'un écran capacitif multi-tactile prometteur, transparent et disposant de meilleurs temps de réponse que les dispositifs précédents. Toutefois, le projet sera abandonné par *Bell Labs* qui ne parvient pas à trouver de débouchés commerciaux pour une technologie encore jeune et méconnue du grand public.

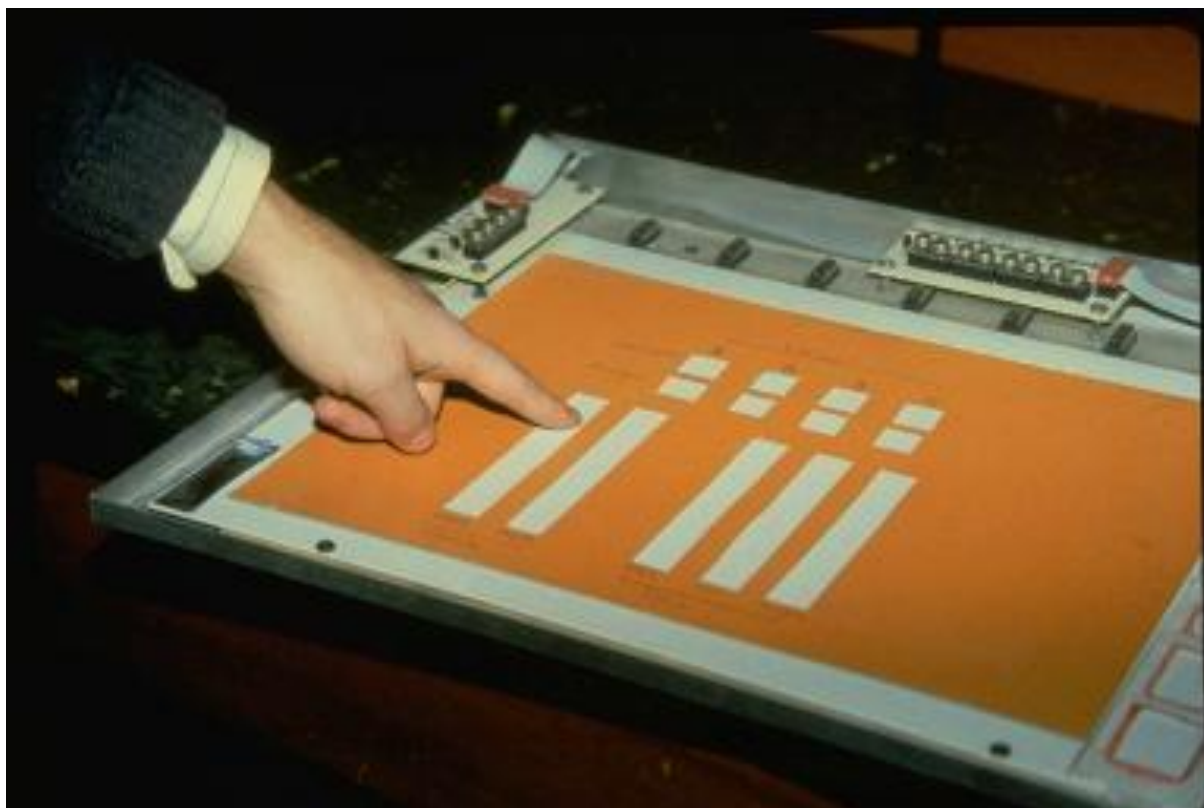


Figure 15 - Tablette multi-tactile capacitive. *Bell Labs*, 1985.

Dans les années 1980, Paul McAvinney, de l'université Carnegie Mellon, travaille sur des systèmes de musique assistée par ordinateur. Il met au point un cadre optique qui se place devant l'écran et permet de distinguer jusqu'à trois doigts simultanément. Il présente ses travaux en **1986**⁸⁸ et monte sa société quelques années plus tard. Le projet initial donnera lieu à la création de *VideoHarp* (cf. Annexes 1.1.2.), un instrument de musique électronique et programmable qu'il conçoit avec Dean Rubine, chercheur en interaction homme-machine dont les contributions sur le sujet ont permis de mettre en évidence des caractéristiques essentielles du geste dans la manipulation d'objets graphiques.

⁸⁸ McAVINNEY Paul. *The Sensor Frame - a gesture-based device for the manipulation of graphic objects*. In *Sensor Frame*, Carnegie-Mellon University, Pittsburgh, 1986

Dean Rubine utilise notamment le *SensorFrame* comme dispositif expérimental lors de sa thèse sur le geste en interaction homme-machine⁸⁹. Le dispositif consiste en une source de lumière qui encadre l'écran et en quatre capteurs optiques, de simples caméras disposées à chaque angle, qui analysent la position des doigts en deux dimensions lorsqu'ils sont à proximité de la surface. L'approche de McAviney est originale et séduit notamment la NASA qui voit la possibilité d'utiliser cette technologie afin d'intervenir sur les équipements scientifiques pour la réparation ou la maintenance. Plus tard, la NASA ajoutera au dispositif la possibilité de déterminer l'angle d'approche de la main lorsque le doigt est près de l'écran.

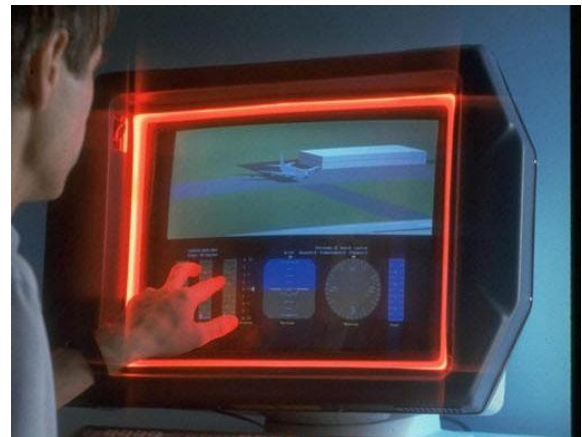


Figure 16 - *SensorFrame*. A gauche, schéma du dispositif, reproduction. A droite, photo du dispositif.
In [McAviney Paul, 1986]

Par la suite, la recherche s'oriente davantage vers les principes de la réalité mixte et donne naissance à des projets comme *DigitalDesk* en **1991**, composé d'un bureau ordinaire surmonté d'une caméra qui capture la surface de travail. Un système analyse les données reçues par la caméra et détecte non seulement la position des doigts mais également le type de documents posés sur le bureau. Un vidéoprojecteur est installé également au-dessus de la surface et peut projeter des images sur le bureau et sur les documents papier posés intentionnellement. Un microphone est installé sous le bureau et permet d'identifier le bruit du doigt tapant sur la surface pour envoyer au système une commande similaire au clic de souris. Il s'agit d'un des premiers et des plus emblématiques dispositifs de réalité mixte qui inspirera notamment Fitzmaurice *et al.* dans leurs travaux, quelques temps plus tard, sur les interfaces tangibles et la notion d'objets palpables.

⁸⁹ RUBINE Dean Harris. *The Automatic Recognition of Gestures*. Thèse. Carnegie Mellon University, Pittsburgh, USA 1991

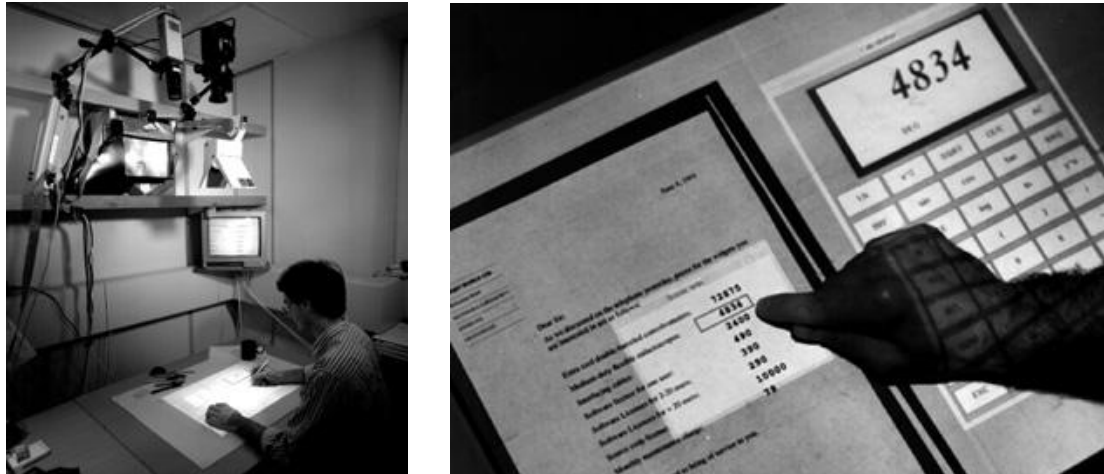


Figure 17 - VideoDesk. A gauche, une vue du dispositif. A droite, un exemple d'interaction.

1998 – 2003 : Premières commercialisations grand public

FingerWorks est une société spécialisée dans la reconnaissance de mouvement, fondée en **1998** par John Elias et Wayne Vesterman. Dès 1999, suite à la thèse de Wayne Vesterman, la société commercialise une série de produits multi-tactiles, comme le « *iGesture Pad* » et le clavier « *TouchStream* », qui a la particularité de s'adresser aux personnes souffrant de stress et de douleurs musculaires. Ces dispositifs sont constitués d'un clavier et de surfaces multi-tactiles qui convertissent les différentes combinaisons et mouvements des doigts en commandes souris ou en macro-commandes. La société dépose de nombreux brevets qui sont rachetés en 2005 par *Apple*.



Figure 18 - A gauche, la tablette iGesture Pad. A droite, le clavier TouchStream. Fingerworks

L'histoire de la technologie multi-tactile se poursuit en **2001**, avec la présentation de la table *DiamondTouch*⁹⁰, une surface tactile multi-utilisateurs, au symposium organisé par l'ACM sur les logiciels et technologies pour les interfaces utilisateur. Tandis que les technologies multi-tactiles antérieures étaient davantage étudiées pour un utilisateur unique, celles-ci présentent la particularité de permettre à plusieurs utilisateurs d'interagir en même temps tout en identifiant quelle est la personne qui est à l'origine du contact. C'est le laboratoire de recherche de Mitsubishi, le MERL (*Mitsubishi Electric Research Labs*) qui est à l'origine de cette technologie et qui l'a principalement développée dans le but de faciliter le travail collaboratif, le brainstorming et la prise de décision lors de réunions de travail autour de documents électroniques. Le dispositif est composé d'une surface tactile capacitive, d'un vidéoprojecteur situé au-dessus de la table et d'une série d'antennes radio (transmetteurs) situées sous la surface. En touchant l'écran, l'utilisateur augmente le signal de certaines antennes vers les récepteurs généralement situés sous la chaise de l'utilisateur et permet au système de savoir qui touche et à quel endroit. Il s'agit d'un des premiers dispositifs multi-tactiles dédié à l'utilisation de collecticiels à affichage unique (*single display groupware*), programmes qui permettent à plusieurs utilisateurs de travailler en collaboration par l'intermédiaire d'un système d'affichage disposant de plusieurs entrées utilisateur. Après une longue phase d'expérimentation et de tests dans différentes universités et laboratoires de recherche, la table de travail multi-tactile du MERL est commercialisée en 2006.

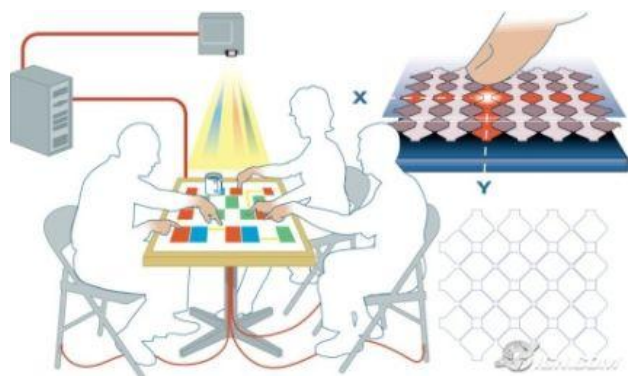


Figure 19 - La table multi-tactile DiamonTouch développée par les laboratoires de *Mitsubishi Electric Research*, 2001.

En **2002**, une compagnie du nom de *JazzMutant* est créée par Guillaume Largillier et Pascal Joguet. La société est principalement orientée vers le développement d'interfaces homme-

⁹⁰ DIETZ Paul et LEIGH Darren. *DiamondTouch : a multi-user touch technology*. In *Proceedings of the 14th annual ACM symposium on User interface software and technology* (UIST '01). ACM, New York, USA, 2001, pp. 219-226

machine pour l'industrie culturelle. Avec l'arrivée de Julien Olivier en 2003, ils développent un écran multi-tactile à technologie résistive, capable d'identifier un nombre illimité de points de contact. Après avoir déposé un premier brevet en 2004, ils sortent un produit appelé *Lemur* en 2005 et le rendent disponible sur le marché des particuliers. *Lemur* est une tablette munie d'un écran LCD de 12 pouces multi-tactile. Il s'agit d'un contrôleur dédié aux applications temps-réel audio et multimédia pour une utilisation musicale, sur scène par exemple. Le dispositif ne fonctionne pas comme un contrôleur midi mais utilise le protocole OSC (*OpenSoundControl*), dont la fiabilité et les performances feront de ce protocole un standard utilisé dans de nombreuses applications multi-tactiles aujourd'hui. Ce contrôleur est fourni avec une large bibliothèque d'objets : boutons, potentiomètres rectilignes ou rotatifs, interrupteurs, claviers, zones d'effets qui permettent à l'utilisateur de contrôler et de générer des sons, des effets sonores ou multimédia en temps réel. Le contrôleur *Lemur* a très certainement inspiré la création de nombreux instruments ou dispositifs d'interaction pour la scène musicale comme la *Reactable* vu précédemment ou le *Tenori-on*, instrument de musique électronique présenté au SIGGRAPH en 2005.



Figure 20 - A gauche, le contrôleur Lemur. A droite le Tenori-on.

2003 – 2007 : laboratoire Microsoft

Microsoft ne fait son apparition sur la scène des technologies multi-tactiles et des interfaces gestuelles qu'à partir de **2003** avec deux projets distincts ; d'une part un dispositif vertical appelé *Touchlight*⁹¹ développé par Andrew Wilson et d'autre part un dispositif horizontal

⁹¹ WILSON Andrew D. *TouchLight: an imaging touch screen and display for gesture-based interaction*. In *Proceedings of the 6th international conference on Multimodal interfaces (ICMI '04)*. ACM, New York, USA, 2004, pp. 69-76

plus largement connu : la table interactive *Surface*. *TouchLight* est composé d'une surface de projection verticale et d'un système de captation optique qui permet le design d'interfaces gestuelles et tactiles. Deux caméras sont situées derrière l'écran translucide et un émetteur infrarouge illumine les mains de l'utilisateur. En s'approchant de l'écran, les mains sont éclairées par l'illuminateur infrarouge et détectées par les caméras qui sont positionnées de part et d'autre de l'écran. La combinaison des deux images permet au système de rendre davantage visibles les objets proches de l'écran et d'occulter le fond.

Dans ce dispositif, l'utilisateur peut interagir avec le système sans forcément être en contact avec l'écran, il peut manipuler l'interface par des gestes devant la surface et valider en « tapant » sur l'écran. C'est un microphone, comme pour *DigitalDesk*, posé sur la surface qui permet de détecter les vibrations émises lorsque l'utilisateur touche l'écran. Depuis 2006, la société *Eon Reality* commercialise une solution d'affichage et d'interaction gestuelle du nom d'*EON TouchLight* qui utilise les procédés mis au point par Microsoft.

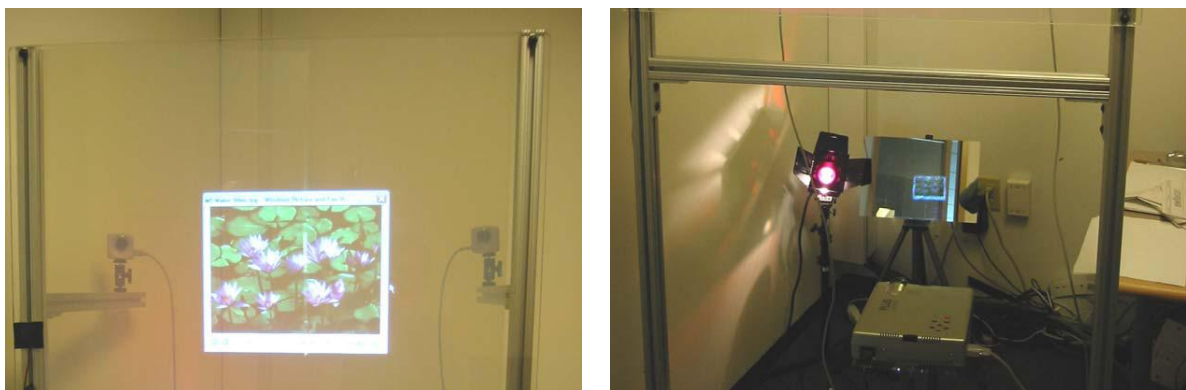


Figure 21 - TouchLight. A gauche, le système de captation. A droite, le système de projection arrière.
In [WILSON Andrew, 2004]

La *Surface* est une table interactive multi-tactile développée par *Microsoft* qui permet à plusieurs utilisateurs de manipuler les données numériques par l'intermédiaire d'une interface multi-tactile gestuelle. L'origine du projet remonte à 2001 où Steven Bathiche et Andy Wilson des laboratoires de recherche de la compagnie abordent le sujet d'une table interactive lors d'une réunion ordinaire. C'est en octobre 2001 qu'une équipe de travail est créée officiellement afin de faire évoluer l'idée de départ et transformer le concept initial en un projet concret. En **2003**, séduit par le concept d'un produit aux interactions riches et intuitives, Bill Gates encourage l'équipe à continuer le développement de leur projet. Quelques mois plus tard, un premier prototype était né, le *T1*, construit sur la base d'une

table *IKEA* et utilisant comme diffuseur du papier vélin d'architecte. Les premières applications développées sont principalement des mini-jeux et un gestionnaire de photos. L'utilisation d'illuminateurs infrarouges, comme pour *TouchLight*, permet aux caméras situées dans la table de « voir » par réflexion de la lumière infrarouge tout objet qui se pose sur la table, qu'il s'agisse de la main et des doigts ou bien de tout autre objet physique. En mettant un marqueur sous la surface de l'objet que l'on pose sur la table, le système est capable d'identifier précisément de quel objet il s'agit (téléphone portable, lunettes, etc.)

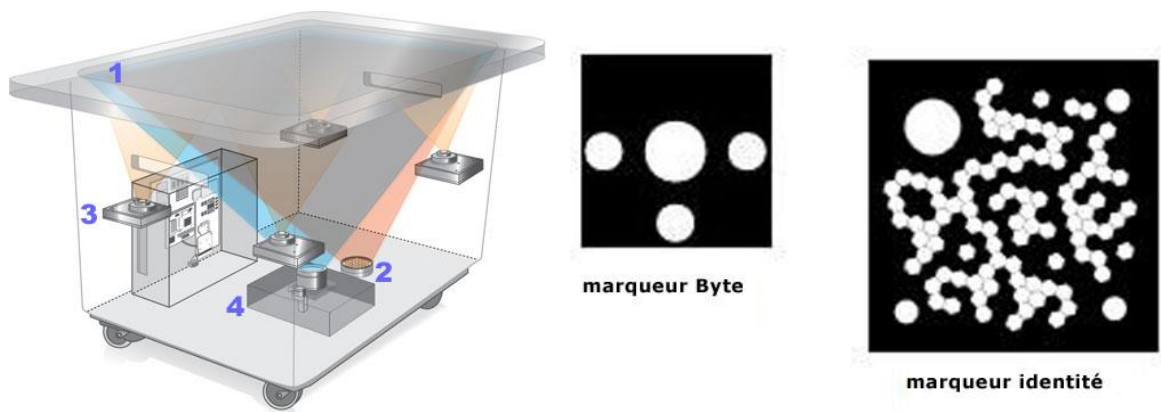


Figure 22 - Microsoft Surface V1. A gauche, un schéma de fonctionnement. 1. L'écran diffuseur 2. L'illumination infrarouge 3. Les caméras IR 4. Le projecteur. A droite, deux types de marqueurs (1 octet et 128 bits) pour identifier les objets.

Les capacités de détection d'objets de la table *Surface* et la quantité importante de points de contact qu'elle est capable d'analyser (une cinquantaine) font de ce dispositif un outil de développement avec un large potentiel. Cependant, les propositions d'interfaces développées jusqu'à présent sur cette surface ne reflètent pas la richesse des interactions promises par la table interactive. En effet, les applications sont généralement commerciales et n'utilisent pas toutes les possibilités offertes par la technologie. L'utilisation de la table par le public suscite chez les designers d'interaction la crainte de créer des interfaces trop complexes et pas suffisamment intuitives car ce mode d'interaction est nouveau pour les utilisateurs. Cela explique en partie pourquoi le pincement (ou « *pinch* »), geste dont l'usage s'est répandu grâce à l'*iPhone* d'*Apple*, est très souvent, dans les applications pour *Surface*, le seul geste qui justifie l'utilisation d'une technologie tactile multi-points.



Figure 23 - A gauche, le prototype T1, 2003. A droite, la Microsoft Surface V1, 2007.

En 2004, le kit de développement et le design définitif sont finalisés en 2005. Il faudra deux ans aux ingénieurs pour passer du prototype à un produit commercialisable (design, industrialisation, kit de développement, documentation, etc.) et voir la *Microsoft Surface* présentée au public en 2007.

2005 – 2010 : Jeff Han et Apple

En 2005, Jeff Han publie de son côté un article⁹² décrivant une technique simple et peu coûteuse pour développer une technologie multi-tactile sur une surface de projection. Reposant sur le principe de la réflexion totale frustrée (cf. Chapitre II - 2.2.), phénomène d'optique géométrique, le système qu'il présente possède plusieurs avantages : la résolution spatiale et temporelle élevée des points de contact, le faible coût du dispositif et la possibilité de le décliner pour des tailles d'écran beaucoup plus importantes. Le prototype développé par Jeff Han est composé d'une surface en polycarbonate transparente dont les bords ont été polis afin d'illuminer la surface par les côtés. Les illuminateurs infrarouges éclairent donc l'intérieur de la surface selon un angle précis et les rayons de lumière ne peuvent s'en échapper à cause du phénomène de réflexion totale frustrée. C'est seulement en posant un objet sur la surface que la lumière s'échappe et vient alors éclairer cet objet. Une caméra équipée de filtres passe-bandes adéquats analyse les objets éclairés par la lumière infrarouge et le système opère les fonctions de traitement d'image standards comme la suppression du fond, un seuil de luminosité et la diminution du bruit vidéo. Au

⁹² HAN Jefferson Y. *Low-cost multi-touch sensing through frustrated total internal reflection*. In *Proceedings of the 18th annual ACM symposium on User interface software and technology (UIST '05)*. ACM, New York, USA, 2005, pp. 115-118

final, l'image après traitement ne montre que les points de contact utilisateur qui sont ensuite identifiés et analysés par le système en fonction de la calibration. Aucun phénomène d'occlusion n'intervient sur la détection (un point de contact ne peut en cacher un autre) et la résolution de la capture dépend essentiellement de la résolution de la ou des caméras utilisées.

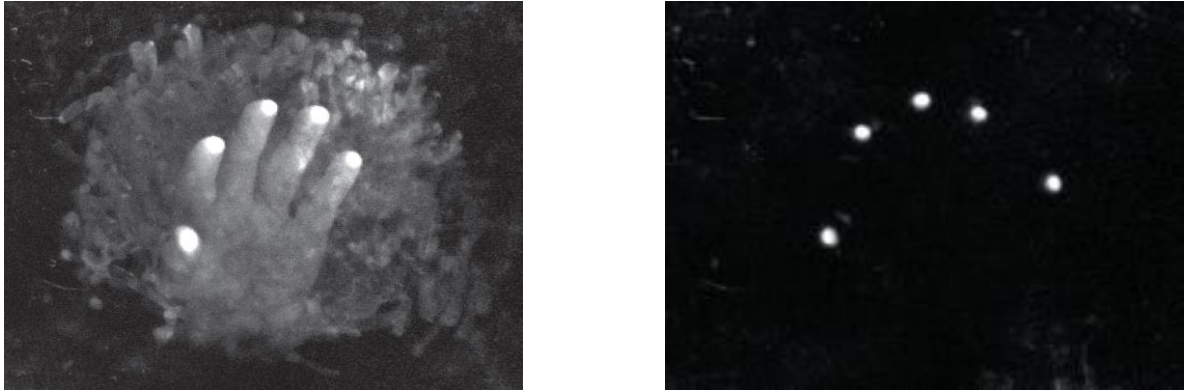


Figure 24 - A gauche, l'image avant le traitement d'image. A droite, l'image après le traitement d'image.

En février **2006**, Jeff Han participe à la conférence TED (*Technology Entertainment Design*) où il présente son prototype et quelques applications d'interactions multi-tactiles, bi-manuelles et originales. La vidéo, diffusée six mois plus tard sur le site d'hébergement de vidéo *Youtube*, connaîtra une large audience puisque la vidéo compte aujourd'hui plus de 350 000 vues. Comme nous le verrons dans le prochain chapitre, il s'agit du premier véritable contact entre le public et la technologie multi-tactile. Suite aux nombreuses sollicitations dont il fait l'objet après son passage à la conférence TED, Jeff Han fonde la société *Perceptive Pixel*, spécialisée dans le développement de technologies tactiles et la recherche sur les interfaces graphiques.

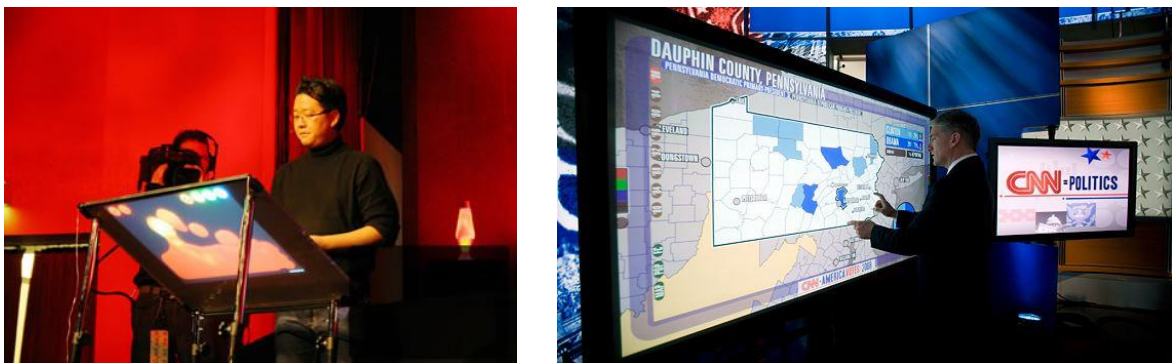


Figure 25 - A gauche, Jeff Han au TED. A droite, le « MagicWall » de Perceptive Pixel utilisé par la chaîne CNN.

Dans l'histoire des écrans multi-tactiles, la sortie de l'*iPhone* en **2007** marque la rencontre définitive entre le public et ces technologies en introduisant une interface simple, tactile et intuitive dont le modèle d'interaction gestuel se répand auprès de millions d'utilisateurs. Sous la direction de Steve Jobs, *Apple* travaillait auparavant sur la mise au point d'une tablette PC (*tablet personal computer*), sorte d'ordinateur personnel équipé d'un écran tactile comme principal périphérique utilisateur. Les premiers travaux sur les tablettes PC ont été menés dès les années 1960 avec notamment l'invention de la tablette *RAND* et la présentation des concepts de l'ordinateur personnel pour enfant, le *dynabook*, inventé par Alan Kay (cf. chapitre I - 1.2.). *Apple* fait sa première entrée dans le marché des tablettes PC en 1989 avec un produit appelé *Newton MessagePad*. Alors qu'initialement, la taille de ce produit devait être plus large, sa taille a été revue à la baisse afin de pouvoir rentrer dans la poche. Il s'agit d'un nouveau type d'ordinateur qui sera appelé assistant numérique personnel ou PDA (*Personal Digital Assistant*).

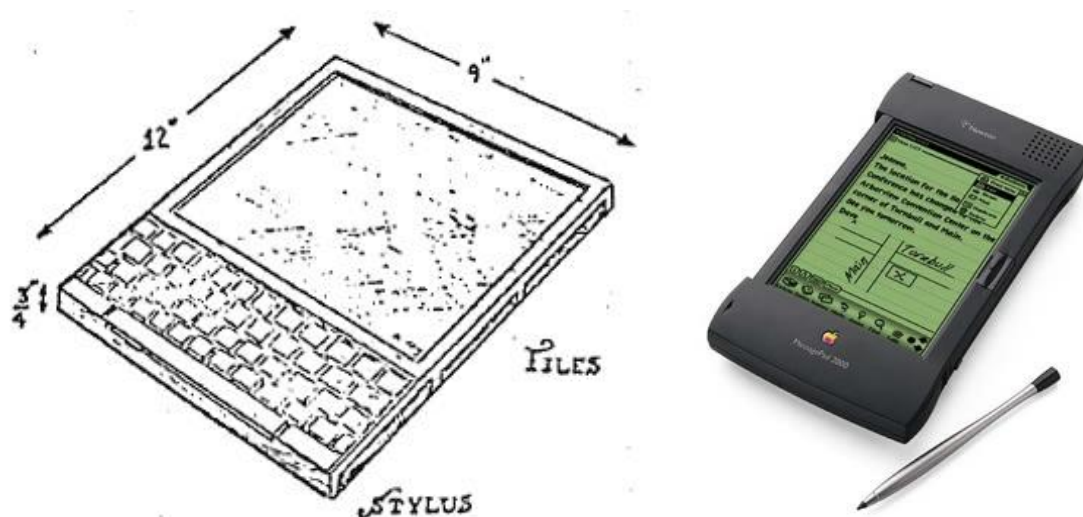


Figure 26 - A gauche, le concept du *Dynabook* d'Alan Kay. A droite, le *Newton MessagePad* d'Apple.

En avril 2003, à la conférence « *D: All Things Digital* » organisée par *The Wall Street Journal*, Steve Jobs fait part de ses convictions sur l'avenir des tablettes PC et il estime qu'il est préférable de renoncer au marché des PDA pour se concentrer sur des dispositifs comme l'*iPod*, un lecteur de médias portable dont la première génération était mise sur le marché en 2001 et dont le contenu pouvait être synchronisé avec un logiciel appelé *iTunes*. Selon Steve Jobs, le téléphone portable est en passe de devenir un dispositif majeur pour accéder et échanger l'information. Après une collaboration avec *Motorola* et un succès mitigé des téléphones *ROKR E1* et *ROKR E2* issus du partenariat entre les deux sociétés, *Apple* stoppe le

développement, n'assure plus le support de ces téléphones et s'apprête à révéler l'*iPhone*, en juin **2007** à la convention *MacWorld*, où l'annonce du smartphone d'Apple attire l'attention non négligeable des médias.

L'*iPhone* est un téléphone mobile disposant des fonctions d'assistant numérique personnel. Il dispose de fonctionnalités telles que l'agenda, la réception de courriers électroniques, la navigation sur Internet, la messagerie instantanée et une puce GPS. Ce qui le démarque des produits concurrents, c'est la présence d'un écran qui couvre quasiment toute la surface de l'appareil et une interface multi-tactile qui propose aux utilisateurs une façon simple et originale de manipuler l'interface par des interactions gestuelles. L'utilisation d'un écran tactile comme dispositif d'interaction n'est pas un phénomène nouveau mais contrairement à la plupart des téléphones ou assistants personnels de même nature, l'*iPhone* dispose d'un écran et d'un système dont les interactions privilégient les contacts avec les doigts plutôt qu'avec un stylet. De plus, la technologie d'écran capacitif permet de manipuler l'interface avec plusieurs doigts en utilisant des gestes imaginés précédemment par la société *FingerWorks*. Suite au succès rencontré par l'*iPhone* auprès du public, *Apple* décide de revenir à la fabrication de tablettes et annonce début 2010 la sortie de l'*iPad* dont la taille de l'écran se situe entre celle d'un *iPhone* et celle d'un ordinateur portable. La tablette dispose des mêmes technologies que celles utilisées pour l'*iPhone* : écran multi-tactile, synchronisation avec iTunes et interactions gestuelles.

Retour en 1991 : Mark Weiser

Les visions de Mark Weiser, considéré comme le père de l'informatique ubiquitaire abordée en préambule de ce chapitre, semblent se réaliser une vingtaine d'années plus tard. En effet, dans un article⁹³ paru en 1991 dans le magazine *Scientific American*, il décrit trois technologies issues du concept d'informatique ubiquitaire (ou ambiante) : les *tabs*, les *pads* et les *boards*. Ce que décrit Mark Weiser peut être compris comme l'informatique sans ordinateurs. Ceux-ci tendent à disparaître au profit de l'électronique et des processeurs miniaturisés qui se répandent dans les objets de notre quotidien. Dans ce contexte, « ubiquitaire » peut s'interpréter comme « partout » mais aussi « en chaque chose ». En

⁹³ WEISER Mark. *The computer for the 21st century*. In *Human-computer interaction*, Ronald M. Baecker, Jonathan Grudin, William A. S. Buxton, and Saul Greenberg. San Francisco, USA, Weiser, pp. 933-940

s'intéressant davantage aux dispositifs imaginés par Mark Weiser et ses collègues de travail, il est surprenant de voir la justesse de leur vision.

« Les tabs sont les plus petits composants de la virtualité incarnée. Parce qu'ils sont interconnectés, les tabs se répandront dans le domaine des ordinateurs miniatures comme les calculettes et les organisateurs de poche. Ils disposeront également de fonctions que ne possèdent pas les ordinateurs aujourd'hui. Par exemple... des badges pourront s'identifier à des récepteurs placés à l'intérieur des bâtiments, cela permettant de suivre la trace des personnes ou des objets auxquels ils sont attachés.⁹⁴ »

Tels que les décrit Mark Weiser, les *tabs* peuvent aujourd'hui être assimilés aux smartphones, portables, qui accompagnent l'utilisateur partout en simplifiant certaines tâches du quotidien et peuvent être géo-localisés grâce à une puce GPS. Les puces RFID entrent également dans cette catégorie car elles permettent de suivre un objet ou une personne en étant détectées par des récepteurs.



Figure 27 - A gauche, le prototype « Tab », 1991. A droite, l'iPhone 4 d'Apple, 2010.

⁹⁴ « Tabs are the smallest components of embodied virtuality. Because they are interconnected, tabs will expand on the usefulness of existing inch-scale computers such as the pocket calculator and the pocket organizer. Tabs will also take on functions that no computer performs today. For example . . . badges can identify themselves to receivers placed throughout a building, thus making it possible to keep track of the people or objects to which they are attached. » In WEISER Mark. *The computer for the 21st century*. In *Human-computer interaction*, Ronald M. Baecker, Jonathan Grudin, William A. S. Buxton, and Saul Greenberg . San Francisco, USA, Weiser, pp. 933-940. Traduction libre : COQUELIN Marine et MORIN Rudy.

En parlant des *pads*, Weiser décrit un dispositif plus large :

« [...] quelque part entre la taille d'une feuille de papier, celle des ordinateurs portables et des ordinateurs de poche. Les pads diffèrent des ordinateurs portables conventionnels de façon notable. Alors que les ordinateurs portables accompagnent leurs propriétaires partout, les pads qui doivent être portés d'un endroit à un autre sont un échec. Les pads sont censés être des ordinateurs fragmentés qui peuvent être pris et utilisés n'importe où ; ils n'ont pas d'identité ni importance propres.⁹⁵ »

Les pads tels qu'ils sont décrits par Mark Weiser nous rappelle étrangement ce que l'on appelle les tablettes comme l'*iPad* d'*Apple*. Connectées à Internet et disposant de fonctionnalités de synchronisation avec les données personnelles des utilisateurs conservées en ligne et accessibles partout, les tablettes font en effet partie d'une informatique fragmentée, dont les informations numériques sont accessibles partout. La synchronisation des données utilisateur, le stockage des données sur serveurs et l'« informatique en nuage » (*cloud computing*), l'accessibilité et la compatibilité des services en ligne pour ces différents dispositifs portables, correspondent à des usages qui entrent fortement en résonance avec les descriptions faites par Mark Weiser.



Figure 28 - A gauche, le prototype « Pad », 1991. A droite, la tablette *iPad* d'*Apple*, 2010.

⁹⁵ « [...] something of a cross between a sheet of paper and current laptop and palmtop computers . . . Pads differ from conventional portable computers in one crucial way. Whereas portable computers go everywhere with their owners, the pad that must be carried from place to place is a failure. Pads are intended to be 'scrap computers' (analogous to scrap paper) that can be grabbed and used anywhere; they have no individualized identity or importance » In *Human-computer interaction*, Ronald M. Baecker, Jonathan Grudin, William A. S. Buxton, and Saul Greenberg . San Francisco, USA, Weiser, pp. 933-940. Traduction libre : COQUELIN Marine et MORIN Rudy.

Plus loin dans l'article, Mark Weiser fait la description des *boards* :

« Dispositifs d'affichage d'un mètre environ (boards) aux usages multiples : à la maison, des écrans vidéo et des tableaux d'affichage. Au bureau, des tableaux d'affichage, des tableaux blancs ou des chevalets de conférence. Une board pourrait aussi servir de bibliothèque électronique à partir de laquelle on pourrait télécharger du texte sur un pad ou un tab⁹⁶. »

La similitude entre les *boards* telles qu'elles sont imaginées par Weiser et la *Microsoft Surface* est troublante. Les fonctionnalités avancées comme la possibilité de télécharger et d'échanger des données entre la table interactive et un smartphone ont déjà été présentées par *Microsoft*. De manière plus générale, les télévisions numériques sont aujourd'hui reliées à des serveurs en ligne, elles permettent pour certaines de naviguer sur Internet (*Google TV* aux Etats-Unis, *Freebox V6* en France) et tout type d'appareil mobile peut se connecter à ce système par l'intermédiaire d'une connexion WIFI.

Aujourd'hui

Le succès de l'*iPhone* et celui de l'*iPad* ont été à l'origine de la rapide et croissante pénétration des écrans tactiles dans les smartphones, lecteurs de médias portables et autres appareils mobiles. Jennifer Colegrove, directrice du département des technologies d'affichage pour le cabinet d'étude *DisplaySearch*, fait un point sur l'évolution du marché des technologies tactiles :

« La pénétration des écrans tactiles a rapidement augmenté dans les téléphones portables, les lecteurs mp3 et médias, les systèmes de navigation, et autres applications. Durant les prochaines années, les écrans tactiles vont connaître

⁹⁶ « Yard-size displays (boards) serve a number of purposes: in the home, video screens and bulletin boards; in the office, bulletin boards, whiteboards or flip charts. A board might also serve as an electronic bookcase from which one might download texts to a pad or tab. » In *Human-computer interaction*, Ronald M. Baecker, Jonathan Grudin, William A. S. Buxton, and Saul Greenberg. San Francisco, USA, Weiser, pp. 933-940. Traduction libre : COQUELIN Marine et MORIN Rudy.

une forte progression pour des applications de taille plus large comme les ordinateurs tout-en-un, les tablettes PC, l'éducation et l'entraînement.⁹⁷ »

Selon *DisplaySearch*, les revenus générés par les écrans tactiles dans les différents secteurs industriels représenteraient 9 milliards de dollars en 2015 contre 3,7 milliards en 2009.

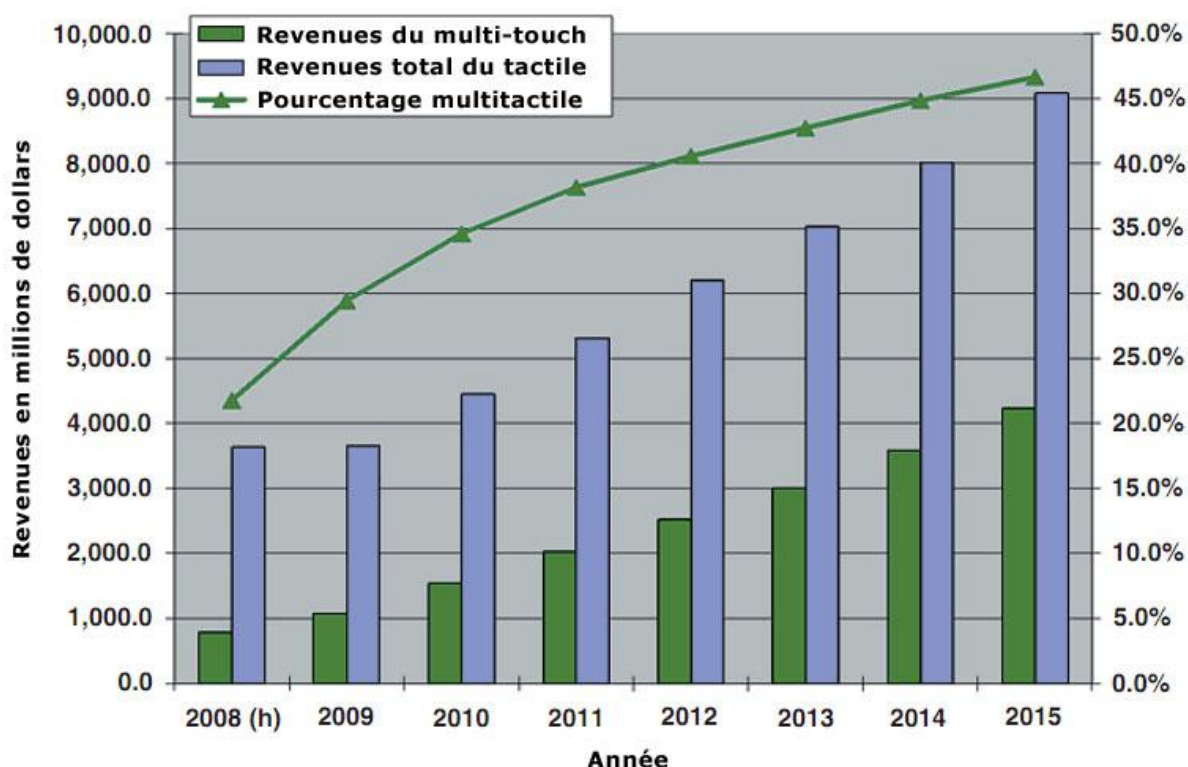


Figure 29 - Prévisionnel des revenus générés par les écrans tactiles. Analyse réalisée par *DysplaySearch*, mars 2010.

Les technologies multi-tactiles ont fortement évolué depuis les années soixante-dix. Toutefois, elles ne semblent pas avoir atteint leur maturité et poursuivent leur évolution. En effet, certaines technologies sont plus encombrantes comme la *Microsoft Surface V1* (illumination infrarouge) tandis que d'autres sont plus discrètes mais ne sont pas aussi précises (cadres optiques infrarouges). Partant de ce constat, Bill Buxton *et al.*⁹⁸ font la démonstration en 2007 d'une technologie multi-tactile prometteuse appelée *ThinSight*,

⁹⁷ « Touch screen penetration has been rapidly increasing in mobile phone, PMP/MP3, portable navigation, and other applications. Over the next several years, touch screens will undergo strong growth in large-size applications such as all-in-one PCs, Mini-note/slate PCs, and education/training » COLEGROVE Jennifer. *The State of the Touch-Screen Market in 2010*. In *Display marketplace*. Mars 2010, pp. 22-24

⁹⁸ IZADI Shahram, HODGES Steve, BUTLER Alex, RRUSTEMI Alban et BUXTON Bill. *ThinSight : Integrated Optical Multi-touch Sensing through Thin Form-factor Displays*. In *Proceedings of the 2007 workshop on Emerging displays technologies: images and beyond: the future of displays and interaction* (EDT '07). ACM, New York, USA, Art. 6

précurseur du principe utilisé dans la table interactive *Microsoft Surface V2* dévoilée en 2011. Elle consiste en une série de capteurs et d'émetteurs optiques infrarouges situés derrière la matrice LCD. Lorsqu'un objet touche l'écran, la lumière infrarouge est réfléchiée et détectée par les récepteurs. Cette technologie a plusieurs avantages. D'une part, étant donné qu'il s'agit d'un système par illumination arrière, le système est capable d'utiliser et de reconnaître des marqueurs. En outre, la précision des capteurs permet d'obtenir des informations riches sur les contacts détectés à l'écran. Il n'existe pas de phénomène d'occlusion comme sur les cadres optiques et il est possible d'analyser la pression exercée sur l'écran.

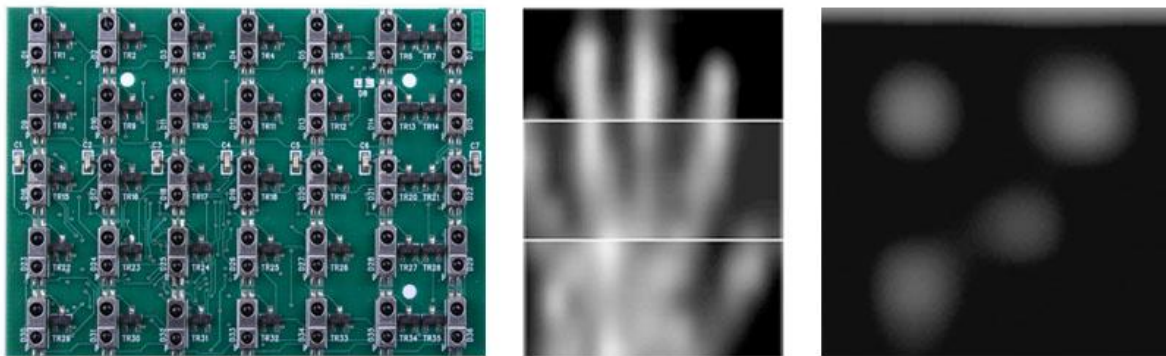


Figure 30 - A gauche, la série de capteurs opto-électronique de la technologie *ThinSight*. Au centre, l'image d'une main capturée par les capteurs. A droite, l'image d'un marqueur posé à la surface de l'écran.

En janvier **2011**, *Microsoft* et *Samsung* annonce *Pixel-sense*, une technologie basée sur *ThinSight*, intégrée dans un écran à cristaux liquides (*liquid crystal display* ou LCD) de 40 pouces qui sert de surface interactive à la deuxième version de la *Microsoft Surface*, connue sous le nom de *SUR40*. Il s'agit d'une couche de capteurs photo-sensibles aux infrarouges qui permet de détecter tout objet posé ou en approche de la surface. Cette modification dans la configuration du dispositif permet de réduire considérablement le volume qui était préalablement nécessaire pour l'installation du vidéoprojecteur, des illuminateurs et des caméras infrarouges. La démonstration de cette technologie multi-tactile réalisée au salon technologique CES 2011, nous donne une idée de la tendance des constructeurs et des directions techniques qui sont en cours. La miniaturisation du dispositif est permise par la réduction des profondeurs d'écrans et l'utilisation de micro-capteurs situés sous la surface. Cela permet d'envisager la modularité de ces écrans, sous la forme de tables interactives adaptées pour les entreprises ou pour la maison, ou d'envisager des écrans larges verticaux dont le volume autorise une installation plus aisée. Mais ce qui attire

d'avantage mon attention, c'est de constater que les chercheurs, plus qu'essayer d'améliorer les performances des dispositifs ou de réduire leur encombrement, augmentent les capacités sensorielles de ces objets. « *Les images savent désormais qu'on les regarde.* »

Les objets de notre quotidien embarquent des technologies invisibles qui leur donnent, de manière anthropomorphique, des sens propres aux humains : le toucher, la vue et l'ouïe. C'est comme si l'informatique ubiquitaire, ou ambiante, au-delà de se réaliser dans notre quotidien, empruntait une voie où les objets se dotent progressivement des mêmes dispositions sensorielles que l'homme. L'image numérique, après s'être parée de la dimension interactive, devient sensible à son environnement, capable de reconnaître des objets, de lire, d'interpréter les gestes, les expressions faciales et la voix de l'homme. Les interfaces tactiles et multimodales suivent des chemins différents mais partagent le même horizon, à la différence que les technologies tactiles « augmentent » les objets de notre quotidien de manière discrète et invisible.



Figure 31 - A gauche, la table interactive SUR40 ou Microsoft Surface V2. A droite, ce que voient les capteurs en niveaux de gris. Sources : <http://www.betanews.com> et <http://blogs.msdn.com>

3. Pré-appropriation technologique

Afin de mieux comprendre les raisons de l'intérêt soudain en faveur des technologies multi-tactiles et les nouvelles propositions d'interface homme-machine, il est nécessaire de revenir sur les facteurs culturels et sociaux qui ont participé à la naissance de la curiosité du public pour ces nouvelles interfaces et technologies gestuelles. La mise au point technologique de ces dispositifs a été progressive, comme nous avons pu le constater dans le précédent chapitre. Pendant toute la période d'incubation des technologies mettant le geste au cœur de l'interactivité, les avancées techniques ne se sont pas uniquement

manifestées lors de conférences scientifiques mais ont traversé d'une certaine manière la couche hermétique du milieu de la recherche pour devenir accessible au plus grand nombre avant même leur commercialisation. Cette forme de mise à disposition des avancées technologiques de la recherche auprès du public a été rendue possible par la multiplication des canaux d'information et par le croisement des domaines scientifiques, ludiques et artistiques.

3.1. L'art numérique

L'influence de l'art contemporain n'est pas négligeable en ce sens où les installations interactives et participatives proposées par les artistes contemporains ont provoqué la rencontre entre le public et des œuvres interactives, gestuelles et tangibles. Dès 1935, le *Space Modulator* de Moholy Nagy permettait au spectateur de modifier la structure de l'œuvre en déplaçant des épingles sur une surface de zinc perforée. Plus tard, les règles du happening, définies par Allan Kaprow, engageaient l'intervention intentionnelle et physique du spectateur dans le processus de l'œuvre. Avec l'arrivée de l'« art par ordinateur » et des premiers univers virtuels, cet art de la participation va produire des œuvres comme l'installation multimédia et multi-sensorielle de Myron Krueger : *VideoPlace* en 1974 (cf. figure 32) ou plus récemment le dispositif expérimental *The Khronos Projector* de Alvaro Cassinelli en 2005 (cf. figure 33). De l'art cinétique à l'art numérique, les rencontres entre le public et les œuvres ou objets interactifs ont contribué, très souvent par leur caractère ludique, à la popularité des interfaces interactives et gestuelles.



Figure 32 - *Videoplace*, Myron Krueger, 1974.



Figure 33 - *Khronos Projector*, Alvaro Cassinelli, 2005.

Les installations numériques créent les conditions d'une expérimentation de l'œuvre. L'activité perceptive du spectateur est conditionnée par l'exploration physique de l'espace et donc par l'expérience sensible. Le corps du spectateur fait l'expérience du « réel » en étant confronté à la matérialité des structures de l'installation, mais il fait également l'expérience du « virtuel » lorsqu'il se saisit de l'interactivité et modifie le cours de l'œuvre. « *L'ordinateur est une machine de Feed-back*⁹⁹ » nous dit Edmond Couchot, ce qui signifie qu'il est capable de rétroagir d'une part sur les actions de l'utilisateur et d'autre part sur la pensée créatrice de celui qui l'utilise. Il est capable de faire surgir chez son utilisateur, par ses propriétés auto-discursives, des idées qu'il n'aurait probablement pas soupçonnées tout seul. Parfois même, les idées ou les formes les plus originales proviennent d'une erreur de manipulation ou de programmation ou bien d'expérimentations liées à l'outil qui alimentent la palette de l'artiste numérique.

L'ordinateur est un outil programmable, c'est-à-dire qu'il est possible de modifier et même de créer de toutes pièces les spécificités propres à son comportement. C'est également un outil capable de créer des outils, Michel Bret parle alors de méta-outil. C'est justement cette particularité qui permet à l'ordinateur d'étendre l'univers des possibles, tant dans la forme de l'œuvre que dans ses processus. Les artistes numériques ont à leur disposition des logiciels de création graphique et de programmation qui leur permettent de réaliser la forme et le comportement de l'œuvre. Toutefois, ces logiciels, programmés par d'autres, répondent déjà à des codes conceptuels qui influencent l'approche artistique et la représentation de l'œuvre. Les artistes numériques sont généralement conscients de cette limite et détournent les fonctions initialement prévues des logiciels ou composent avec plusieurs applications afin de s'affranchir de ces contraintes. À l'inverse, les artistes peuvent trouver de l'aide auprès d'un ingénieur afin d'exploiter pleinement la liberté permise par l'ordinateur.

Il existe plusieurs logiciels à vocation artistique, souvent gratuits et open-sources, c'est-à-dire dont le code source est librement distribué, qui sont destinés aux artistes, aux chorégraphes et aux musiciens. Il s'agit généralement d'interfaces de type programmation visuelle (*Visual Programming Language* ou VPL) basées sur l'assemblage de « briques » de fonctions paramétrables qui, une fois interconnectées, constituent un programme ou un

⁹⁹ COUCHOT Edmond et HILAIRE Norbert. *L'art numérique*. Flammarion, Paris, 2005, 265 p. (coll. Art)

algorithme de captation et de rendu graphique. L'ensemble des briques de développement disponibles permet de prototyper rapidement et simplement un dispositif de départ, ce qui en fait un outil particulièrement apprécié des artistes. Parmi eux, je citerai des logiciels dont j'ai fait l'expérience et ont tous la particularité de pouvoir analyser le geste ou traiter des données issues de capteurs :

- *MaxMsp* est la combinaison du logiciel MAX (*IRCAM/Cycling'74*), pour le contrôle temps réel d'applications musicales et multimédia interactives,
- *PureData* est la version open-source de *MaxMsp*,
- *Processing* est issu d'une volonté de simplifier le langage de programmation afin de le rendre plus accessibles par les artistes,
- *Openframeworks* est un kit de développement en C++ spécialisé dans le traitement et la production d'image en temps réel,
- *VVVV* est un logiciel de création multimédia,
- *EyesWeb*.

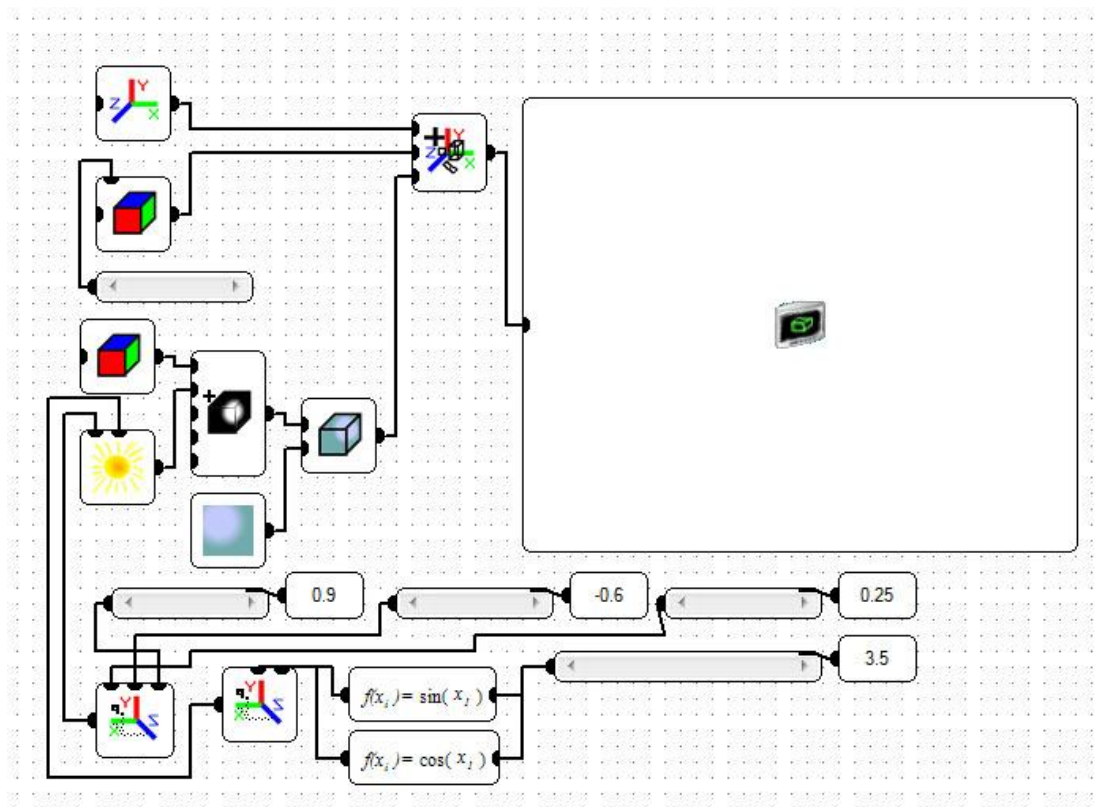


Figure 34 - Exemple de programmation graphique réalisé avec EyesWeb 5.0.1.4

Je ferai un court développement sur le logiciel eyesweb pour avoir été pendant plusieurs années en contact avec le laboratoire qui est à l'origine du projet. *Infomus Lab* est un laboratoire de recherche universitaire qui a été fondé en 1984 à l'université de Gênes, et dont les chercheurs (mathématiciens et informaticiens) mènent leur recherche et développement (*Eyesweb*, *Harp*, *VideoBrain*) sur des interfaces multimodales interactives dédiées aux arts numériques (installations, performances, danses). Ils orientent leur axe de recherche vers l'acquisition et la compréhension d'expressions gestuelles non verbales, en vue d'obtenir des modèles d'analyse qui permettraient de saisir l'émotion de l'interacteur. Leurs études portent sur l'acquisition et l'interprétation du mouvement, sa segmentation et ses caractéristiques. Ils s'appuient notamment sur les recherches en psychologie du mouvement de Boone et Cunningham qui se sont penchés sur la signification émotionnelle des mouvements du corps en vue de décoder certaines émotions comme le bonheur, la tristesse ou la colère. Ils s'intéressent également aux théories du mouvement de Rudolf Laban et sur les outils d'analyse de la spatialité, des qualités temporelles ou énergétiques du geste qu'il a mis au point.

3.2. Les univers de science-fiction

Ce sont également les univers futuristes de la littérature de science-fiction (Philip K. Dick, Arthur C. Clarkes, Isaac Asimov, etc.), souvent adaptés en images pour le cinéma, qui augmentent l'aura technologique de ces dispositifs et provoquent la rencontre anticipée entre le public et ces technologies. En 1968, Arthur C. Clarke publie « *2001 : L'Odyssée de l'Espace* », scénario du film éponyme dont le travail d'écriture a été mené avec le réalisateur Stanley Kubrick, où apparaissent de nombreuses technologies futuristes dont l'écran de contrôle (*Control Screen*) ou le *NewsPad*, un ordinateur de la taille d'un portable et dédié à la lecture des informations et de textes. Clarke et Kubrick décrivent avec une extrême justesse, avec quasiment quarante ans d'avance, comment l'homme utilise aujourd'hui les dispositifs portables, smartphones et tablettes.

Ils font en effet la description d'un système d'exploitation dont le fonctionnement nous rappelle étrangement celui de l'iOS d'Apple que l'on trouve sur l'iPhone ou l'iPad :

« Lorsqu'il en eut assez des rapports officiels et des mémos, il brancha son minibloc d'information sur le circuit du vaisseau et parcourut les dernières nouvelles de la Terre. Il formait l'un après l'autre les numéros de code des principaux journaux électroniques du monde. Il en connaissait par cœur la plupart et n'avait pas besoin de consulter la liste qui figurait au dos du bloc. En jouant sur la mémoire de la visionneuse, il pouvait consulter la première page et choisir rapidement les rubriques qui l'intéressaient. Chacune avait son propre numéro de référence et, lorsqu'il le formait, le rectangle qui avait les dimensions d'un timbre-poste s'agrandissait sur l'écran. La lecture achevée, il suffisait de revenir à la vision de la page entière et de choisir une autre rubrique.¹⁰⁰ »

En 1982, une table interactive est aperçue dans le film *Tron* de Steven Lisberger. Il s'agit d'une table tactile bi-manuelle et en réseau puisqu'elle permet de se connecter au Maître contrôle principal (*Master Control Program*), un ancien programme d'échecs dont l'intelligence artificielle a évolué. Dans le même temps, Nimish Mehta, étudiant à l'université de Toronto, travaille sur l'interface machine flexible. Le film montre au monde entier une table interactive alors que l'idée commence à peine à germer dans le milieu de la recherche. Depuis, bien d'autres films ont dépeint les possibilités et les futurs probables des technologies multi-tactiles, tangibles ou gestuelles comme *Minority Report* de Steven Spielberg en 2002, *The Island* de Michael Bay en 2005, *Quantum of Solace* de Marc Forster en 2008 et *District 9* de Neill Blomkamp en 2009. Scott Derrickson fait même apparaître la table interactive *Microsoft Surface* dans un film qu'il réalise en 2008 : *The Day the Earth Stood Still*. Ces blockbusters, étant largement diffusés, alimentent l'imaginaire des téléspectateurs et participent au phénomène d'adoption technologique du public et des futurs utilisateurs.

¹⁰⁰ CLARKE Arthur C. 2001 : *L'Odyssée de l'espace*. Traduction : DEMUTH Michel. J'ai lu, 2001, 190 p.

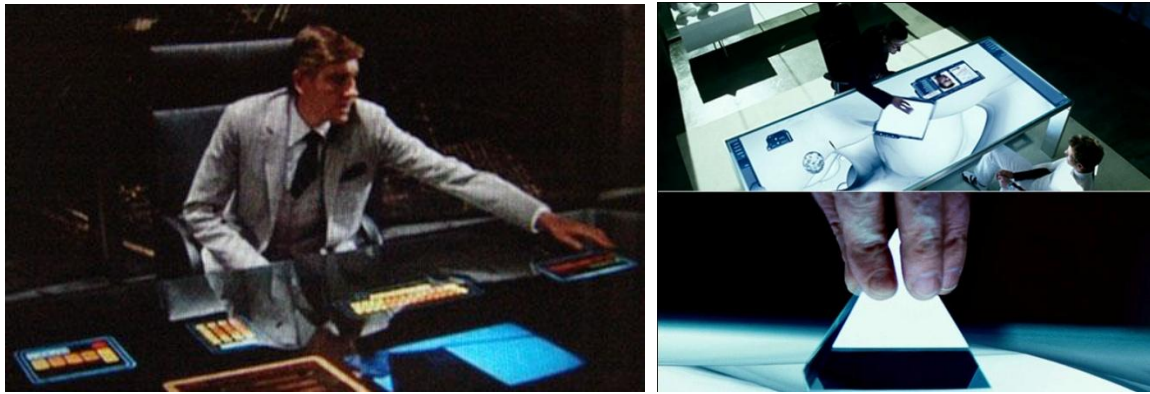


Figure 35 - A gauche, *Tron*, 1982. A droite, *The Island*, 2005.



Figure 36 - *Minority Report*, Steven Spielberg, 2002.

Ces phénomènes de correspondance ne sont pas isolés puisque l'histoire des sciences et des techniques foisonne d'exemples similaires. Nous pourrions nous interroger sur l'origine de ces croisements entre art, science et littérature en questionnant l'idée même d'imaginaire collectif. D'une part, à l'image de Léonard de Vinci, bien des artistes sont également philosophes, écrivains, mathématiciens ou ingénieurs. Par exemple, l'écrivain Arthur C. Clarke n'était pas simplement l'auteur des œuvres de science-fiction qu'on lui connaît, il rédigeait aussi des articles scientifiques remarquables comme celui qu'il publie en 1945 sous le titre de « *relai extraterrestre*¹⁰¹ » dans le magazine *Wireless World*, où il fait la proposition et la description scientifique précise du concept de satellite géostationnaire qui ne deviendra réalité qu'une vingtaine d'années plus tard.

¹⁰¹ CLARKE Arthur C. *Extra-Terrestrial Relays published*. In issue of *Wireless World*, novembre 1945

D'autre part, l'homme se construit un imaginaire à partir de fragments appartenant à un répertoire commun de formes, de symboles, de textes et d'images qui alimentent son imagination et ses idées. Ces fragments, composés d'éléments issus de la perception, participent à notre vision de l'avenir ou à la compréhension de notre monde contemporain. C'est en piochant dans ces fragments que les écrivains, les artistes et les scientifiques s'inspirent mutuellement pour « inventer » le présent.

« [...] le meilleur usage que l'on puisse faire de la science-fiction aujourd'hui est d'explorer la réalité contemporaine au lieu d'essayer de prédire l'avenir... La meilleure chose à faire avec la science aujourd'hui, c'est de l'utiliser pour explorer le présent.¹⁰² »

Par cette remarque, l'écrivain de science-fiction William Gibson illustre comment nous pouvons appréhender certains domaines scientifiques comme la robotique en lisant les textes d'Isaac Asimov ou en visionnant l'œuvre majeure de Ridley Scott : « *Blade Runner* », adaptée du roman « *Les androïdes rêvent-ils de moutons électriques* » de Philip K. Dick; comment le livre « *Ubik* », écrit en 1966 par le même auteur, nous apporte des éléments de réflexion sur les mondes virtuels. Et à l'inverse, comment nous pouvons interpréter le « temps technologique » dans lequel nous vivons à travers le prisme de la science.

En appliquant ce mode de lecture aux croisements entre l'art et la science, nous pouvons constater combien les nombres et les mathématiques ont pu inspirer la peinture, depuis les figures répétitives de l'Antiquité jusqu'aux aberrations géométriques d'Escher ou combien le principe de combinatoire est profondément présent dans la « bibliothèque de Babel » de Jorge Luis Borges ; comment les théories déterministes de Charles Darwin et les sciences naturelles ont inspiré le mouvement naturaliste en littérature ; comment la théorie de la relativité a pu inspirer les textes de Jean-Paul Sartre (« *l'être et le néant* ») ; comment Marcel Duchamp, passionné par la représentation du monde formulée par le mathématicien et philosophe Henri Poincaré et reprise par les auteurs Lewis Carroll¹⁰³ (« *Dynamique d'une*

¹⁰² « [...] best use today is the exploration of contemporary reality rather than any attempt to predict where we are going... The best thing you can do with science today is use it to explore the present. » In GIBSON William. Interview, CNNfn, 26 août 1997

¹⁰³ Lewis Carroll, de son vrai nom Charles Lutwidge Dodgson. Mathématicien et romancier.

particule », 1865) et Herbert George Wells (« *La machine à explorer le temps* », 1895), a désiré faire de la quatrième dimension un concept central de ces œuvres.

3.3. Les communautés de chercheurs

Le phénomène de pré-appropriation des interfaces post-WIMP est permis également par la publication sur Internet des avancées technologiques issues de la recherche scientifique dans les domaines des IHM, de la vision par ordinateur et de l'ergonomie qui sont à l'origine de la communication grandissante autour de ces technologies. En 2006, Jeff Han, chercheur au sein du *Courant Institute of Mathematical Sciences* de l'Université de New York, fait la démonstration public d'un écran translucide multi-tactile et provoque l'enthousiasme général à la conférence TED 2006 (*Technology Entertainment Design*) en montrant non seulement une technologie particulièrement fiable et rapide mais également un ensemble de propositions de design interactif et de modèles d'interactions innovants et dédiés à ce type de dispositifs.

La figure 37 présente un graphique issu de *Google Trends*, outil issu de *Google Labs* qui permet de connaître la fréquence à laquelle un terme a été recherché dans le moteur de recherche Google. En tapant le terme anglais « *multi-touch* » (multi-tactile), le graphique nous montre la popularité du mot clé exprimé sur l'axe des ordonnées (en fréquence relative) en fonction des années. En mettant en parallèle les innovations récentes dans le domaine des technologies multi-tactiles et le graphique exporté par *Google Trends*, nous pouvons facilement reconnaître que la conférence TED animée par Jeff Han est en grande partie à l'origine de la prise de contact entre le public (les internautes) et cette technologie. Nous pouvons également attribuer l'accroissement constant de la curiosité des internautes à la communication excessive des industriels impliqués dans l'utilisation de cette technologie.

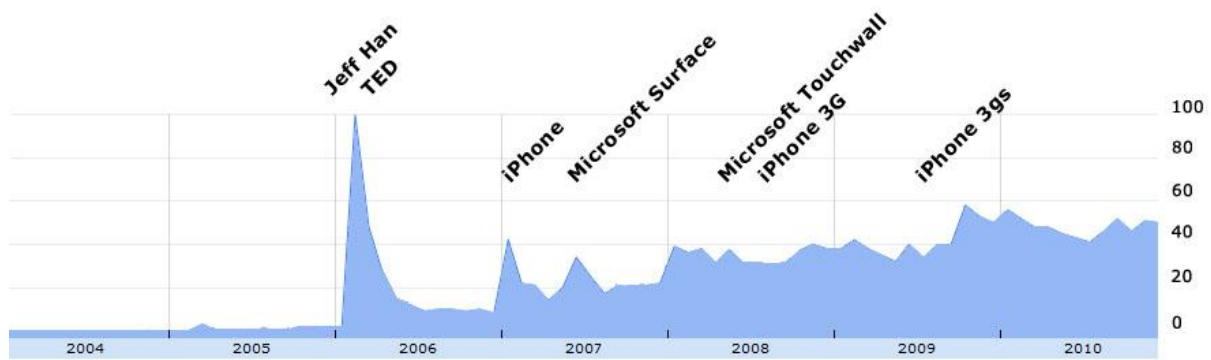


Figure 37 - Evolution du nombre de recherches (en valeur relative par rapport au nombre maximum) du terme "multi-touch" depuis 2004 sur le moteur de recherche *google* mis en perspective avec l'annonce de technologies relatives au multi-tactile.

L'impatience et la curiosité provoquées par ces nouveaux usages en matière d'interface ont rapidement réuni les internautes sous forme de communautés (blog, forum, vidéo en ligne, etc.), partageant une idée simple : créer, utiliser et détourner, pour un coût relativement faible, ces technologies interactives. Ainsi pouvons-nous trouver, sur des sites spécialisés ou sur *diy.com*, *make.com* ou *instructable.com*, les instructions complètes pour réaliser des dispositifs interactifs, afin d'en reproduire le fonctionnement ou au contraire d'en élargir le spectre des usages.

L'engouement pour la création et la fabrication « amateur » de ce type de dispositifs, qui répond à des besoins et à des intérêts multiples, a donné lieu à de multiples détournements technologiques. C'est le cas de la *Wiimote*, bardée de multiples capteurs de précision, accéléromètres et autres périphériques infrarouges, dont le faible coût en a fait un instrument privilégié pour les étudiants, les chercheurs et les amateurs de technologies interactives dans la réalisation et l'expérimentation d'interfaces tangibles ou gestuelles. Johnny Chung Lee, doctorant en interaction homme-machine à l'université de Carnegie Mellon, a su par exemple utiliser la *Wiimote* et la barre capteur (*sensor bar*) de la Wii afin de créer différents dispositifs interactifs comme une interface gestuelle, un tableau multi-tactile ou encore un environnement immersif en 3D dont les effets de perspective réagissent à la position de l'utilisateur. Les tutoriaux et les vidéos de Johnny Chung Lee, disponibles entre autres sur *Engadget*, *Ars Technica*, *YouTube* et *Dailymotion*, ont provoqué un tel enthousiasme que cela a poussé son auteur à monter le forum *wiimoteproject.com*, lieu d'échange de projets, d'idées, de conseils, de programmes amateurs, d'exemples, de vidéos

et de toute autre ressource autour de la fabrication de dispositifs interactifs utilisant la *Wiimote* comme principal capteur.

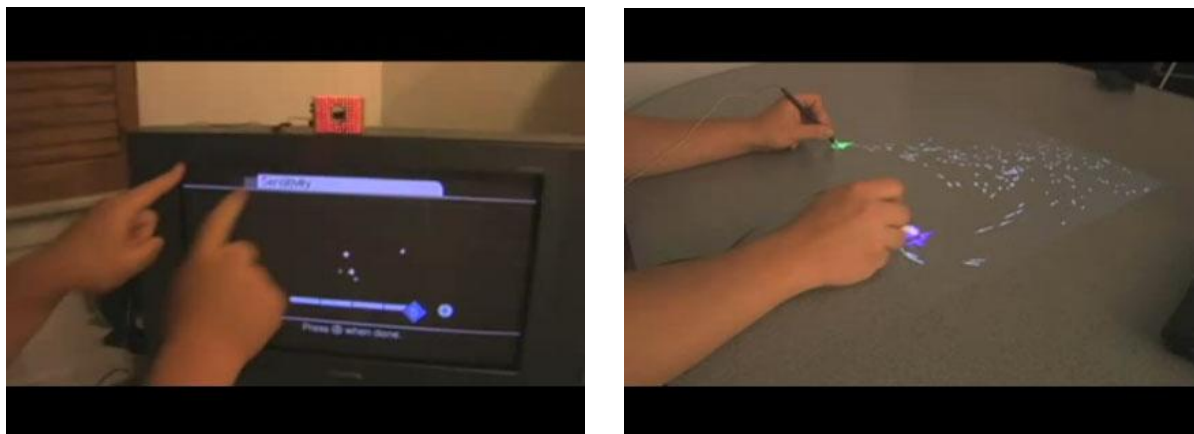


Figure 38 - A gauche, un dispositif de captation gestuelle. A droite, une surface multi-tactile. Les capteurs utilisés pour ces deux projets sont ceux de la *Wiimote*.

Malgré l'apparent amateurisme de ces réalisations, elles participent et concourent à la création de projets plus ambitieux. C'est le cas du pôle d'intérêt *Natural User Interface*, dont l'ambition est de réunir amateurs, chercheurs et artistes autour d'intérêts communs comme de nouvelles méthodes et concepts en interaction homme-machine. De ce projet est né par exemple *Touchlib*, une solution logicielle très répandue pour l'acquisition, l'analyse et le traitement de données issues de tables multi-tactiles. Le projet s'est même associé au « *Google Summer of Code* » dès 2008 afin de voir les réflexions et les expérimentations aboutir à la création de solutions logicielles open-sources. Depuis, le projet *Touchlib* est devenu *Community Core Vision* ou *CCV*, une solution open source et multiplateformes que nous utilisons à Virtual-IT dans le cadre des développements sur les prototypes de surfaces tactiles.

3.4. Le monde vidéo-ludique

Avec la sortie de la console de Nintendo, la *Wii*, la pratique du jeu vidéo se dote d'une nouvelle dimension, en réduisant l'écart symbolique entre l'espace réel et l'espace simulé. Le joueur évolue et interagit dans un espace intermédiaire contraint à la fois par la matérialité de son corps et de son environnement physique mais également par les limites et les règles du jeu qui régissent l'environnement simulé. Le gameplay, et plus précisément le « play » dans le cas présent, c'est-à-dire le jeu comme liberté d'action et de découverte, enrichit l'expérience ludique et ouvre également la voie à la création de nouveaux territoires

vidéo-ludiques. C'est justement cette fenêtre d'expérimentation et de création qui a suscité chez les joueurs l'envie d'explorer ou de détourner des situations de jeux par l'intermédiaire des manettes sensorielles de la *Wii*.

Parmi les projets de « détournement » de l'usage initial des manettes de la *Wii*, nous pouvons évoquer *GlovePie* (*Glove Programmable Input Emulator*), un logiciel qui permet de jouer et de contrôler n'importe quel logiciel sur ordinateur par l'intermédiaire de dispositifs comme des joysticks, des contrôleurs MIDI (*Musical Instrument Digital Interface*), des gants de données (*datagloves*) et des *Wiimotes*. Ce logiciel donne la possibilité aux joueurs de revisiter le gameplay de l'ensemble de leurs jeux PC, voire de mener l'expérience plus loin en créant des scripts d'interactions spécifiques pour tel ou tel jeu. C'est ainsi qu'est né le projet amateur de créer un mode de jeu spécialement élaboré pour *Half Life 2*, le célèbre titre de *Sierra Entertainment*, qui tente d'amener l'expérience de jeu sur le plan de l'immersion physique. L'intérêt principal pour les joueurs est de revisiter un jeu vidéo, selon de nouvelles règles d'interaction, un nouveau gameplay, une nouvelle « expérience » de jeu, afin d'élargir le mode exploratoire des univers virtuels.

Enfin, c'est toute la sphère du jeu vidéo qui est concernée et notamment la création de jeux en ligne avec le logiciel *Adobe Flash*. Dans ce domaine, nous pouvons citer le projet *WiiFlash*, créé par Joa Ebert et Thibault Imbert « pour le plaisir » selon ses auteurs, qui offre aux développeurs *Flash* un moyen simple d'utiliser la *Wiimote* dans leurs projets. L'originalité du gameplay proposée par ce type de jeu a largement contribué à la démocratisation du détournement de la *Wiimote* chez les joueurs mais également de manière plus large, le phénomène a touché l'art numérique et les chercheurs en IHM. La sortie de la technologie de capture de mouvement, la *Kinect*, proposée récemment par *Microsoft*, semble emprunter les mêmes chemins qui la détourne de son usage initial pour profiter aux artistes et aux communautés amateurs puisque quelques semaines après sa sortie, des drivers non officiels existaient déjà afin de la connecter à un système d'exploitation comme *Linux* ou *Windows*.

CHAPITRE III ESPACE DE CONCEPTION

« [...] la main semble bien être non pas un outil, mais plusieurs. Car elle est pour ainsi dire un outil qui tient lieu des autres. C'est donc à l'être capable d'acquérir le plus grand nombre de techniques que la nature a donné l'outil de loin le plus utile, la main. »

Aristote

Au cours des chapitres précédents, nous avons fait le constat de la richesse et de la diversité des contributions scientifiques et industrielles dans le domaine des technologies multi-tactiles gestuelles. Au regard des différentes orientations conceptuelles et techniques de ces technologies, il apparaît qu'il existe plusieurs techniques de détection du toucher et d'acquisition du geste. En fonction des configurations et des caractéristiques techniques utilisées dans la capture du geste, les données issues de leur traitement ne sont pas toujours identiques d'un système à l'autre. Ces technologies se manifestent sous de multiples formes que je souhaite présenter dans ce chapitre afin de proposer une lecture plus complète de mon objet d'étude. Afin de révéler les différents usages, technologies et configurations de ces dispositifs, je propose d'établir une taxinomie des interfaces tactiles gestuelles.

Cette approche s'inspire de la classification proposée par Maria Karam et M.C. Schraefel¹⁰⁴ qui ont établi une classification des interactions homme-machine gestuelles selon quatre catégories : le domaine d'application (*application domain*), le périphérique d'entrée (*enabling technologie*), les périphériques de sortie (*sortie*) et la gestuelle (*gesture style*).

« Avec cette approche, nous espérons que les chercheurs et les développeurs d'interactions gestuelles pourront commencer à profiter des connaissances issues des différents domaines afin de voir les interactions gestuelles sortir des laboratoires de recherche et gagner les applications informatiques de tous les jours.¹⁰⁵ »

Le diagramme ci-dessous (cf. figure 39) nous présente une classification générale des technologies gestuelles. En nous attachant au domaine exclusif des technologies multi-tactiles, nous souhaitons placer notre analyse dans le prolongement des travaux engagés par Maria Karam et M.C. Schraefel, en adaptant et en précisant les différentes catégories de la classification qu'ils proposent.

¹⁰⁴ KARAM Maria et Schraefel M.C. *A taxonomy of gestures in human computer interaction*. In *Transactions on Computer-Human Interactions*, TBD, 2005

¹⁰⁵ « With this approach, we hope that interaction researchers and developers of gesture based interactions can begin to build on the knowledge gained across the different domains to move gesture based interactions out of the research labs and into everyday computing applicationsA taxonomy of Gestures in Human Computer Interaction. » In *ibid*.

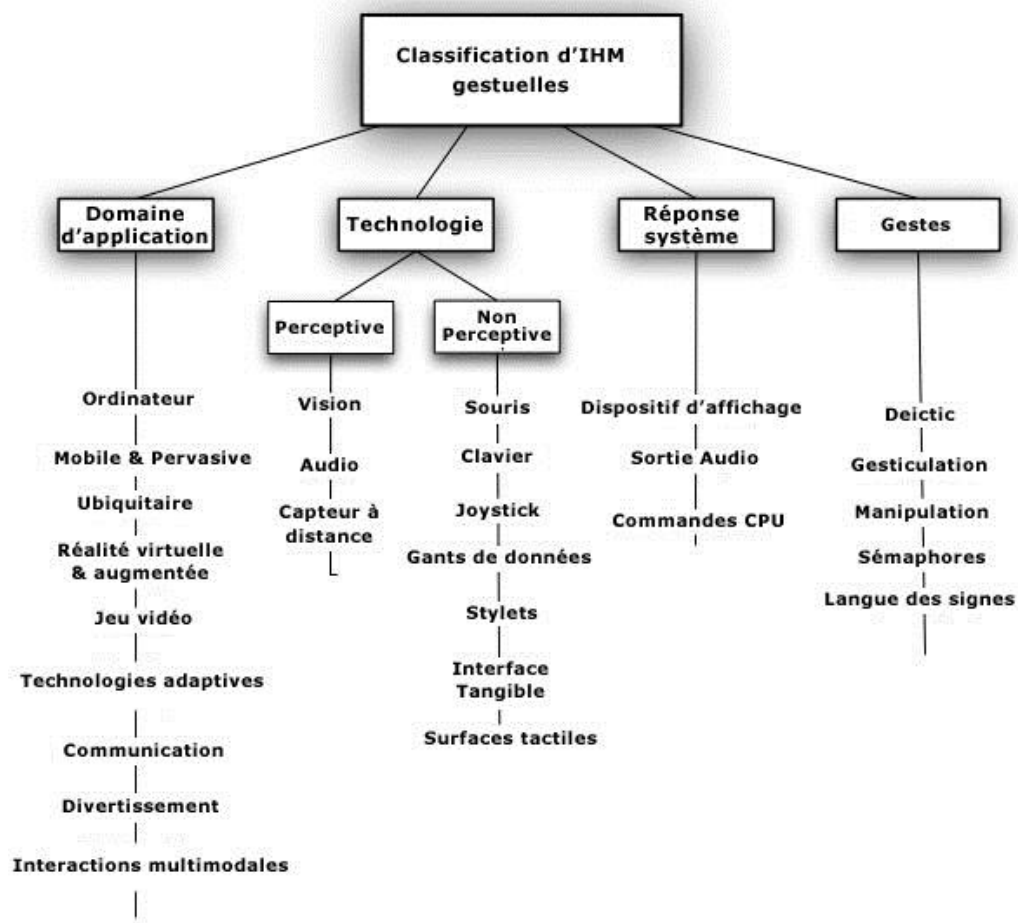


Figure 39 - Classification des interactions homme-machine gestuelle. Le diagramme montre l'organisation de la recherche sur les gestes selon quatre catégories utilisées dans la taxinomie de Maria Karam et M.C. Schraefel. Reproduction. Traduction libre.

1. Domaines d'application

Nous pouvons identifier trois types de marchés distincts couverts par les technologies tactiles. Un marché commun, celui des loisirs comme le jeu vidéo ou la musique qui représente, pour l'année 2009, le secteur générant le plus de revenus¹⁰⁶ grâce notamment à la téléphonie mobile (48,2%), aux lecteurs de médias portables (7,1%) et aux systèmes portables d'aide à la navigation (7,3%). Un marché vertical où la technologie répond essentiellement à des besoins de communication, de vente ou d'applications dédiés. Et enfin, le marché des technologies de l'information et de la communication qui concerne principalement le secteur de l'informatique mobile et des ordinateurs personnels. Dans le

¹⁰⁶ COLEGROVE Jennifer, HSIEH Calvin. *Touch Panel Market Analysis Report*. DisplaySearch, 2010

tableau ci-dessous figurent les principaux domaines d'application des technologies multi-tactiles.

MARCHÉ	DOMAINE	DISPOSITIFS	REVENUS
LOISIRS	Téléphonie	Smartphone	66 %
	Livres numériques (Ebooks)	Lecteurs portables / tablettes	
	Médias (audio, vidéo, podcasts)	Lecteur média portable / tablettes	
	Jeu vidéo	Consoles de jeu portable	
	Navigation	Système d'aide à la navigation	
	Photographie	Appareil photo / imprimante	
MARCHÉ VERTICAL	Communication / Vente	Bornes / Ecrans / Tables	26,6 %
	Point d'information (musée, lieux publics)	Bornes / Ecrans	
	Transports, Banques	Bornes de retrait	
	Civil / Militaire	Systèmes portables et mur d'images	
	Casino	Machine à sous	
	Education / santé	Tables interactives / Ecrans tactiles	
TIC	Ordinateur tout-en-un	Ecrans multi-tactile	7,1 %
	Ordinateur	Ecrans multi-tactile	
	Professionnel	Assistant personnel	
	Domotique	Ecrans de contrôle	
	Accessoires informatiques	Pavé tactile / souris	

Tableau 1 - Domaines d'application des technologies multi-tactiles et répartition des revenus par secteur.
[COLEGROVE, HSIEH, 2009]

2. Périphériques d'entrée multi-tactiles

Au cours de ces trois années de recherche en collaboration avec l'entreprise Virtual-IT, et dans l'optique de créer un environnement de développement (*framework*) compatible avec de nombreuses technologies tactiles, nous avons été contraints d'aller plus loin dans la compréhension de leur fonctionnement. D'une part, parce qu'elles définissent un périmètre de données propre au type de capteurs utilisé et d'autre part, parce qu'elles disposent de possibilités et de contraintes techniques différentes qui ont des conséquences non négligeables dans la qualité du design d'interaction. En effet, comme chaque dispositif technique multi-tactile possède ses propres mécanismes pour localiser les points de contact

des doigts sur la surface et déterminer leur trajectoire, il n'existe pas pour le moment une interprétation unique et conventionnelle des gestes. De plus, une meilleure compréhension de ces dispositifs techniques a été indispensable dès lors que nous avons décidé de construire nos propres prototypes. C'est pourquoi nous faisons l'inventaire des technologies multi-tactiles les plus répandues en mettant en exergue les avantages et les inconvénients théoriques de chacune et la qualité (précision, vitesse, périmètre) des données tactiles et gestuelles qu'elles sont en mesure de capturer.

2.1. Technologies non perceptives

Technologie capacitive

Cette technique fut à l'origine développée pour des interactions tactiles simples (un contact à la fois). Malgré un coût généralement plus élevé que les surfaces résistives, les surfaces capacitives disposent de qualités qui leur permettent de répondre à de multiples usages. C'est une technique largement utilisée dans l'industrie des smartphones, des consoles de jeu portables et des points de retrait ou d'information publics. D'une part, contrairement aux surfaces résistives, le haut degré de transparence de la couche tactile ne nuit pas à la clarté de l'écran situé dessous. D'autre part, les surfaces capacitives projetées autorisent l'ajout d'un verre de protection qui permet de prévenir l'écran des chocs occasionnels et de l'usure liée aux nombreuses sollicitations tactiles qui sont réalisées dans le cas d'une installation permanente. Cela en fait un candidat idéal pour l'installation de dispositifs publics ou pour les appareils mobiles.

Dans ce type de systèmes, la couche tactile accumule les charges électriques et lorsque l'on touche la surface, certaines charges sont transférées aux doigts. Le déficit énergétique est alors détecté par des capteurs, situés aux quatre coins de l'écran, qui envoient cette information à un microprocesseur capable d'interpoler les coordonnées du ou des points de contact. La précision des données captées est généralement proportionnellement inverse aux nombres de points de contact appliqués à la surface. Dans le cas de la technologie capacitive projetée, c'est en mesurant l'état électrique d'une grille de capteurs très fine, située entre deux verres de protection, qu'il est possible de détecter la capacitance générée par le doigt. Les écrans capacitifs peuvent être utilisés par tout type d'interacteurs

conductibles et ne sont pas forcément limités aux contacts des doigts. En fonction du contrôleur, des capteurs et du firmware installé, ils sont capables de gérer des interactions multi-tactiles, de manière limitée toutefois quant au nombre de contacts simultanés.

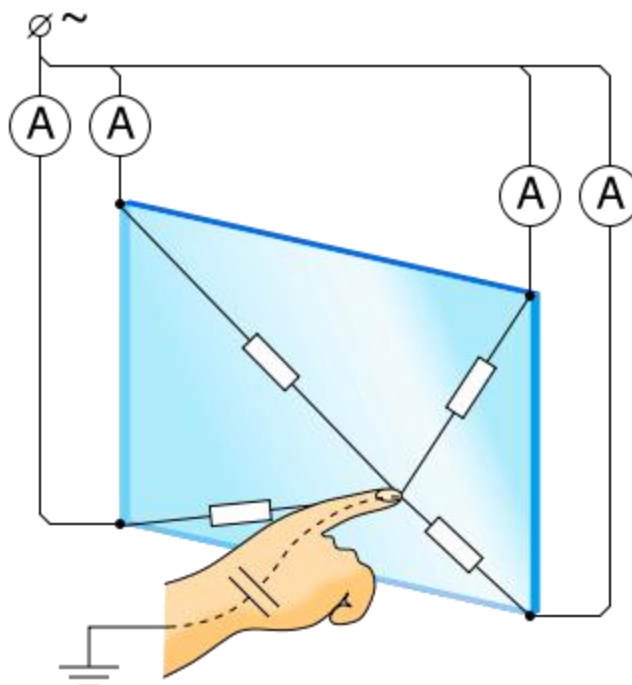


Figure 40 - Illustration simplifiée du fonctionnement d'un écran capacitif.

Cette technologie est pour l'heure limitée par la taille de la surface à couvrir. Lorsque les dalles capacitives excèdent la taille de 20 pouces, cela a pour effet de dégrader rapidement la qualité, les temps de réponse et la précision du système. De plus, le nombre de points de contact étant limité, elle ne convient donc pas pour des écrans larges dont la taille induit la manipulation bi-manuelle et parfois multi-utilisateurs.

Technologie résistive

Tout comme les technologies capacitives, les technologies résistives étaient initialement à point de contact unique. Cependant, certains constructeurs comme Fujitsu, *Touchco*, *HTC* et la société française *Stantum*, à l'origine notamment du projet *JazzMutant*, ont fait évoluer ce dispositif pour des interactions multi-points. Un des avantages de cette technologie tient à son faible coût de production, qui a eu pour conséquence la diffusion plus large de cette technologie dans les produits grand public. Jusqu'à présent, les écrans résistifs bénéficiaient d'une assez mauvaise réputation à cause du manque de clarté de l'image et de la faible

sensibilité tactile de la dalle, parfois déconcertante lors de l'utilisation. Cependant, les efforts engagés dans cette voie technologique montrent aujourd'hui des résultats plutôt convaincants. En effet, la configuration de ce type d'écran autorise l'utilisation des doigts et d'objets comme le stylet contrairement aux écrans capacitifs, la matrice tactile est capable de s'auto-calibrer, la consommation électrique est revue à la baisse et le taux de rafraîchissement des données en entrée a nettement augmenté depuis les premières générations. De plus, la surface tactile est sensible à la pression des doigts et dans certains cas, elle n'est quasiment pas limitée en nombre de points de contacts effectués simultanément.

Un écran tactile résistif est généralement constitué de deux couches conductrices électriquement séparées par un petit écart. Lorsqu'un doigt ou un objet est posé sur la surface, les deux couches métalliques entrent en contact à un point précis, déterminé par le microcontrôleur qui détecte un changement de l'état électrique.

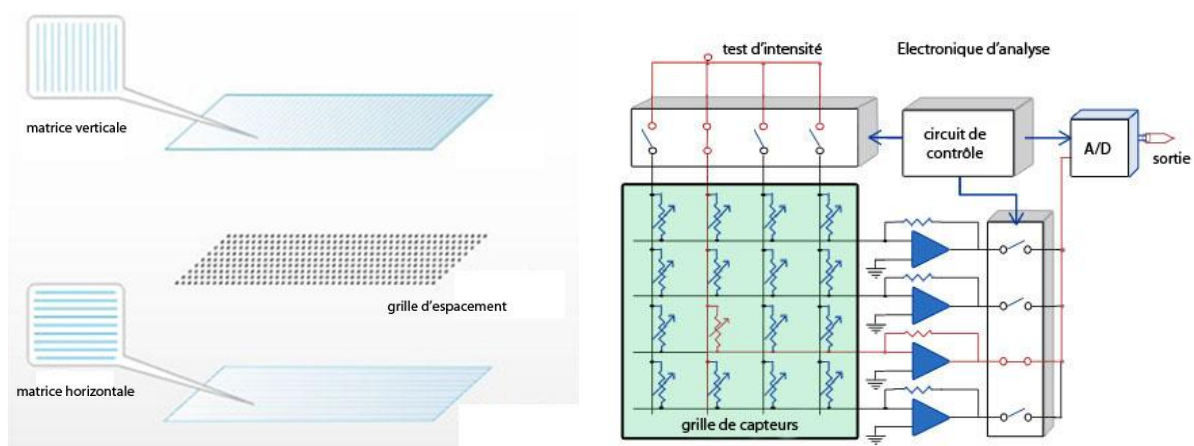


Figure 41 - A gauche, schéma des matrices conductrices d'un écran résistif. A droite, le circuit électrique. Sources : <http://www.tekscan.com> - <http://www.touchuserinterface.com>

2.2. Technologies perceptives

Qu'il s'agisse de technologies tactiles par capteur optique ou par caméra, ces techniques partagent les mêmes principes de capture et de traitement d'image. La quasi-totalité de ces systèmes repose sur l'emploi de la lumière infrarouge et implique dans certains cas des conditions d'éclairage maîtrisées afin de ne pas perturber la détection par des interférences lumineuses. La simplicité et le faible coût de ce type de configurations optiques rendent possible la création de prototypes d'écrans multi-tactiles à des fins d'expérimentation. De

plus, la nature de ces dispositifs offre une grande flexibilité de taille d'écran et de configurations possibles comme la projection avant, arrière ou l'utilisation d'un écran plat LCD ou PLASMA. Dans cette partie, nous ne couvrirons que les configurations qui ont retenu notre attention lorsque nous avons commencé l'élaboration de prototypes multi-tactile. Il ne s'agit donc pas d'une liste exhaustive des technologies optiques existantes mais de la présentation des plus courantes, en donnant un aperçu de leurs avantages, de leurs contraintes et de la simplicité de leur mise en œuvre.

Spectre lumineux

Quasiment toutes les technologies optiques multi-tactiles reposent sur des principes construits sur les bases du traitement d'image ou du signal généré par des capteurs sensibles à la lumière infrarouge. L'illumination est réalisée à l'aide de diodes électroluminescentes ou lasers émettant dans une partie du spectre lumineux du proche infrarouge ou PIR (entre 780nm et 900nm). Les avantages sont nombreux. D'une part, les diodes électroluminescentes dissipent peu d'énergie, donc peu de chaleur. D'autre part, elles sont capables d'émettre dans une bande passante de lumière étroite, facilement identifiable sur le spectre lumineux par un capteur. Dans le cas des systèmes basés sur la configuration par réflexion totale frustrée ou par illumination diffuse, l'acquisition et le traitement d'image se font dans le proche infrarouge afin de créer une image ou un signal de référence par une caméra ou un capteur sensible à cette bande passante. Dans le cas des installations de type projection, les caméras utilisées sont munies d'un filtre passe-bande afin de n'être sensibles qu'aux objets éclairés par les illuminateurs. Cela permet par exemple de ne pas subir l'interférence que pourrait créer la lumière émise par le vidéoprojecteur puisqu'elle se situe dans les longueurs d'onde du spectre visible.

Cependant, ce type de configuration est généralement sensible à la lumière ambiante car la lumière du soleil, comme celle générée par les lampes à incandescence (halogènes, ampoules traditionnelles, etc), émettent sur la quasi-totalité du spectre lumineux et peuvent interférer sur la détection des capteurs ou de la caméra. D'une part, les variations de lumière (passage d'un nuage par exemple) augmentent ou diminuent l'intensité globale d'illumination infrarouge qui ne correspond plus à la référence d'intensité lumineuse de départ. Dans ce cas, il est possible de corriger cela en mettant à jour régulièrement (selon un taux de rafraîchissement) cette référence. D'autre part, l'augmentation de l'illumination

dans le spectre infrarouge par une source parasite réduit l'effet des sources lumineuses du dispositif et tend à réduire la brillance des points de contact qui sont alors plus difficiles à identifier. Une solution à ce problème est illustrée par Johannes Schöning *et al.*¹⁰⁷ et consiste à utiliser l'illumination pulsée. Contrairement à un éclairage constant, les diodes sont allumées par pulsations très courtes (d'une centaine de microsecondes à quelques millisecondes) augmentant ainsi l'intensité de la lumière émise. Ces pulsations sont synchronisées avec la caméra et permettent d'augmenter également le contraste des points de contact.

Réflexion totale frustrée

Le principe d'optique connu sous le nom de réflexion totale frustrée (*Frustrated Total Internal Reflection* ou FTIR) décrit le comportement des ondes électromagnétiques transmises à l'intérieur d'un matériau et ne pouvant s'en échapper car elles sont entièrement réfléchies par les bords. C'est le principe qui est utilisé pour transmettre de l'information, sous la forme de lumière, à travers une fibre optique. Ce phénomène intervient du moment que l'indice de réfraction du matériau est supérieur à celui qui l'entoure et que l'angle d'incidence de la lumière est supérieur à une valeur critique. Dans le cas de la configuration d'une surface multi-tactile telle que l'a présenté Jeff Han¹⁰⁸ en 2005, une série de diodes électroluminescentes (LED) illuminent les tranches d'une surface en polycarbonate. La lumière infrarouge la traverse en rebondissant sur les bords sans jamais les traverser sauf lors du contact d'un doigt qui modifie localement l'indice de réfraction à la surface et permet à la lumière de s'échapper à cet endroit. En s'échappant, la lumière illumine la pulpe du doigt qui peut être à ce moment précis détectée par une caméra, sensible aux infrarouges et située derrière la surface. Une série de filtres et d'algorithmes de traitement d'image permet ensuite de « nettoyer » l'image capturée et de déterminer la position des doigts touchant la surface. Généralement, une couche intermédiaire (en silicone par exemple) est ajoutée sur la surface en polycarbonate afin d'augmenter la différence d'indice de réfraction entre l'intérieur et l'extérieur et réduire ainsi la pression nécessaire à exercer sur la surface afin d'obtenir l'effet d'illumination FTIR.

¹⁰⁷ SCHÖNING Johannes et al. In *Technical Report TUM-I0833, Technical Reports of the Technical, University of Munich*, 2008

¹⁰⁸ HAN Jefferson Y. *Low-cost multi-touch sensing through frustrated total internal reflection*. In *Proceedings of the 18th annual ACM symposium on User interface software and technology (UIST '05)*. ACM, New York, USA, 2005, pp. 115-118

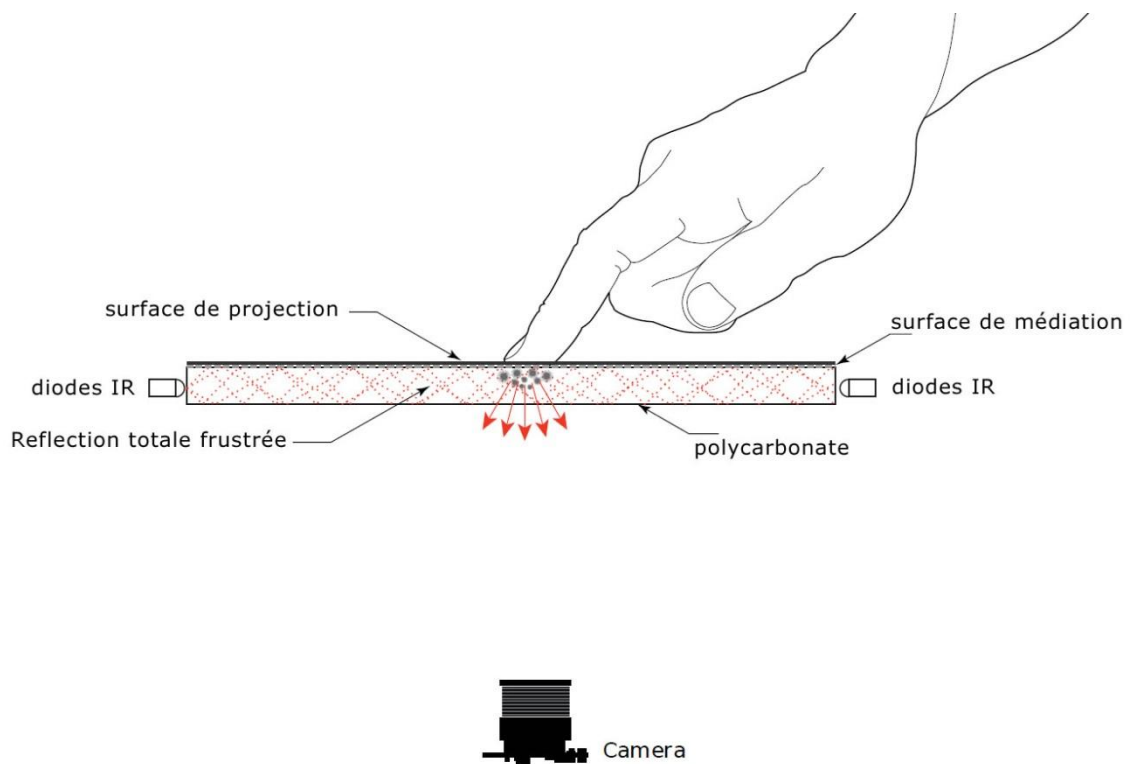


Figure 42 - Configuration d'un système FTIR. Reproduction. In [SCHÖNING, 2008].

L'usage de ce type de surface, appelée surface de médiation (*compliant surface*), permet non seulement de générer des points de contact sans pression mais il bloque également une grande partie de la pollution lumineuse donnant à l'image capturée un contraste plus important et donc plus simple à traiter. Enfin, selon la matière utilisée, cette surface permet de créer un système sensible à la pression des doigts afin d'enrichir les données issues de l'interaction.

Illumination diffuse

Les principes mis en œuvre dans les systèmes d'illumination diffuse fonctionnent de manière similaire aux systèmes FTIR car la détection du point de contact est rendue possible par l'illumination de la pulpe du doigt par de la lumière infrarouge. Dans la configuration d'illumination diffuse (*Diffused Illumination* ou DI), les émetteurs infrarouges ne sont pas situés sur la tranche mais à l'arrière de la surface en polycarbonate. De cette manière, les objets situés devant réfléchissent la lumière en direction de la caméra qui peut alors les détecter.

Contrairement aux systèmes FTIR, les objets situés à proximité de la surface sont également illuminés par la lumière infrarouge dont l'intensité décroît en fonction de la distance. Ce type de système fonctionne grâce à l'utilisation d'une surface diffusante et non transparente. Cette spécificité offre l'avantage d'utiliser la reconnaissance d'objets, par leur forme ou par un marqueur identifié par le système. Enfin, un verre de protection peut être placé devant la surface étant donné que la détection des objets est traitée par un seuil de luminosité qui définit la sensibilité du contact. L'inconvénient de cette configuration tient au fait qu'il est difficile d'obtenir une illumination homogène de la totalité de la surface, d'autant plus s'il s'agit d'un écran de grande dimension. Dans de nombreux cas d'utilisation, en fonction de la sensibilité de la caméra et de nombreux facteurs liés au traitement d'image, les contacts sont parfois détectés avant même que l'utilisateur ait touché physiquement l'écran, ce qui peut créer des points parasites ou aberrants et perturber la reconnaissance par le système.

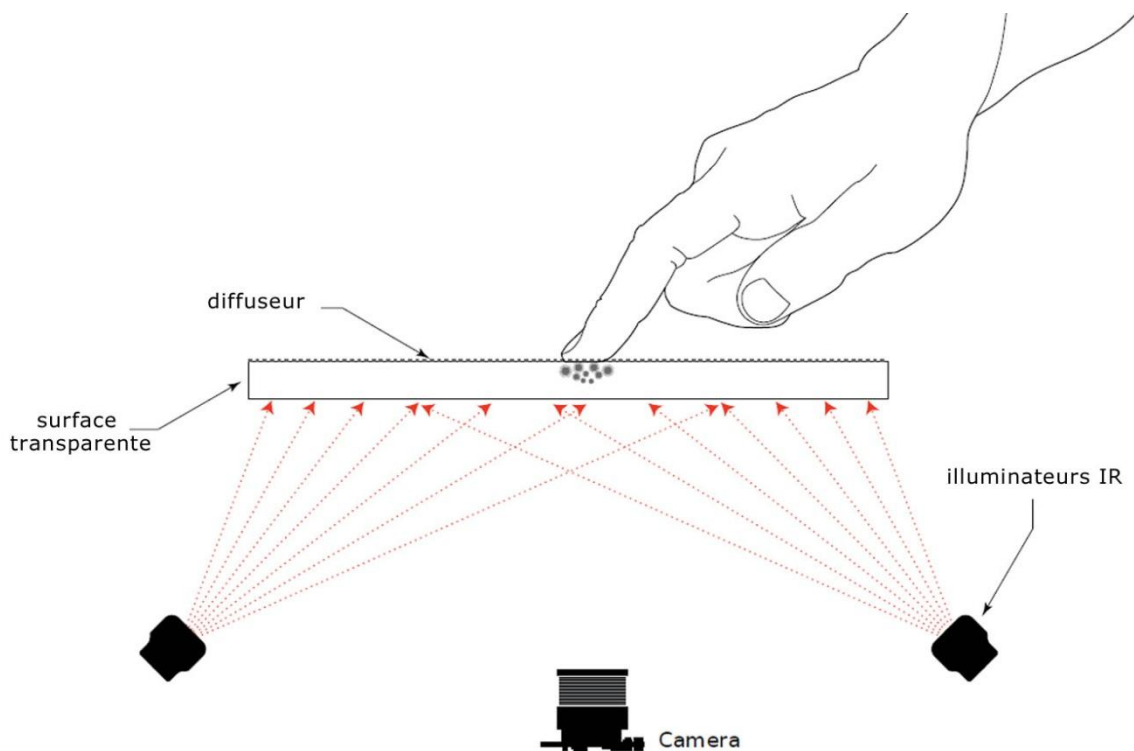


Figure 43 - Configuration d'un système DI. Reproduction. In [SCHÖNING, 2008].

La configuration de type DI peut prendre deux formes différentes : l'illumination arrière, vue précédemment mais aussi l'illumination avant. Cette configuration ne nécessite pas obligatoirement le recours à des illuminateurs infrarouges. La lumière ambiante, issue du soleil ou de lumières halogènes par exemple, émet suffisamment de lumière infrarouge pour permettre au système de fonctionner. Pour la détection, le principe est alors inversé, la caméra ne détecte pas les points de lumières infrarouges mais au contraire les ombres que provoquent les doigts lorsqu'ils sont posés sur la surface diffusante.

Illumination de surface diffuse (DSI)

Cette méthode utilise une surface en polycarbonate spéciale afin de diffuser la lumière infrarouge. La configuration s'inspire de celle utilisant le principe d'illumination par réflexion totale frustrée, sauf que dans ce cas précis, la lumière infrarouge traverse le matériau en rencontrant sur son chemin des milliers de petites particules agissant chacune comme un petit miroir. Lorsque l'on éclaire la tranche de la surface, la lumière infrarouge est redirigée et diffusée sur l'ensemble de la surface en polycarbonate. En remplaçant la lumière infrarouge par de la lumière visible, nous verrions la surface s'illuminer comme si elle était auto-luminescente. L'effet est alors semblable à la configuration DI, la lumière diffusée par la surface éclaire les objets situés à proximité et permet à la caméra de les repérer.

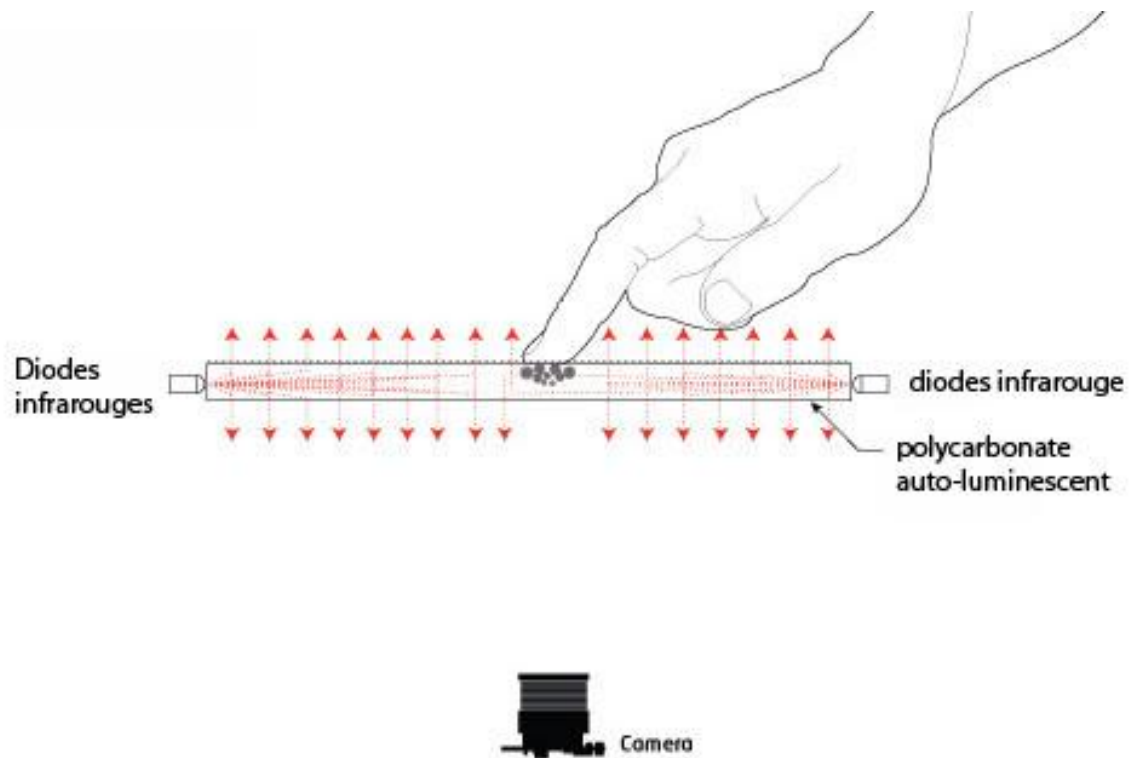


Figure 44 - Configuration d'un système DSI. Reproduction. In [SCHÖNING, 2008].

Plan de lumière laser infrarouge

Dans ce type de dispositifs, un plan de lumière infrarouge (*Laser Light Plane* ou LLP) est émis quelques millimètres devant la surface. Ce plan de lumière est généré par des modules laser de faible intensité munis d'une lentille génératrice de ligne. Selon les configurations, ce plan de lumière laser peut varier de 1 à plusieurs millimètres. Lorsqu'un doigt est posé sur la surface, il va traverser le plan de lumière et une partie de la pulpe du doigt est alors éclairé

par une vive lumière infrarouge. La détection et le traitement d'image est semblable aux dispositifs de type DI, les points de lumière infrarouges étant identifiés par le système de captation.

Selon la taille de l'écran et la puissance des lasers, il peut être nécessaire d'utiliser plusieurs sources lumineuses. La plupart des dispositifs utilisant cette méthode sont équipés de deux à quatre lasers, situés aux angles de l'écran. Les tests que nous avons effectués sur cette technique nous ont également montré l'intérêt d'utiliser plusieurs lasers afin d'éviter les problèmes d'occlusion qui pourraient survenir, c'est-à-dire qu'un doigt puisse en cacher un autre, empêchant ainsi la lumière infrarouge d'illuminer la totalité des doigts posés sur la surface.

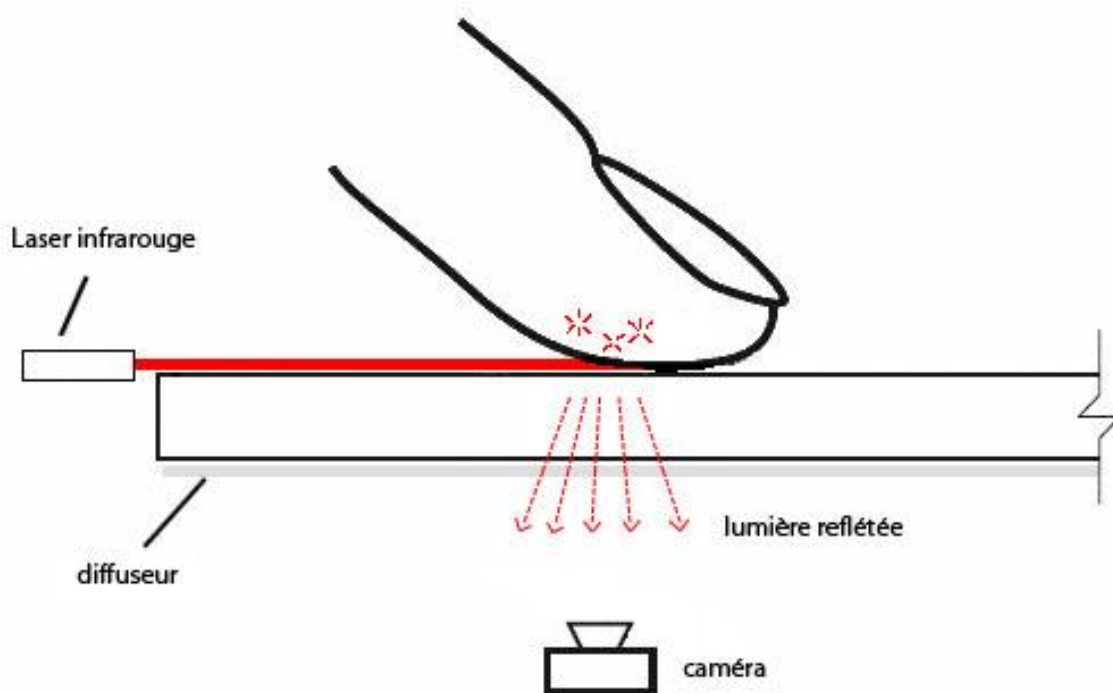


Figure 45 - Configuration d'un système LLP. Reproduction. In [SCHÖNING, 2008].

2.3. Technologies à capteurs optiques

Il s'agit de cadres que l'on place devant un écran de type LCD, PLASMA ou devant un écran de projection. Peu épais, ils constituent une solution pratique et peu onéreuse pour transformer n'importe quelle surface en périphérique multi-tactile. Les technologies actuelles permettent de supporter des tailles d'écran jusqu'à 200 pouces et un nombre considérable de points de contact (une trentaine pour certains modèles).

Ce dispositif est constitué d'un cadre dont les côtés sont munis de bandes infrarouges auto-luminescentes (fibre optique à côté diffusant par exemple) et de capteurs optiques de grande précision situés généralement dans les angles du cadre. Un verre de protection est également régulièrement présent dans ce type de configuration afin de pouvoir adapter ce système à tout type d'écran. Les bandes lumineuses infrarouges créent une fine couche de lumière, juste devant la surface, qui est dirigée vers les capteurs optiques, créant un signal de référence uniforme. Lorsqu'un objet traverse le faisceau, les capteurs détectent une baisse d'intensité lumineuse venant d'une direction et d'une position qui peut être déterminée par des techniques de triangulation.

Cette configuration offre l'avantage de s'adapter facilement à tous types de besoin et peut fonctionner avec les doigts comme avec un quelconque objet. Le système, de par sa nature, peut fournir des informations de contact limitées. Il est capable de renseigner la position, l'orientation et la taille des objets posés sur l'écran, mais ne peut par exemple déterminer avec précision la pression exercée sur un point de contact. La détection de marqueurs n'est pas non plus possible étant donné que les capteurs optiques sont situés sur les côtés et non à l'arrière de la surface.

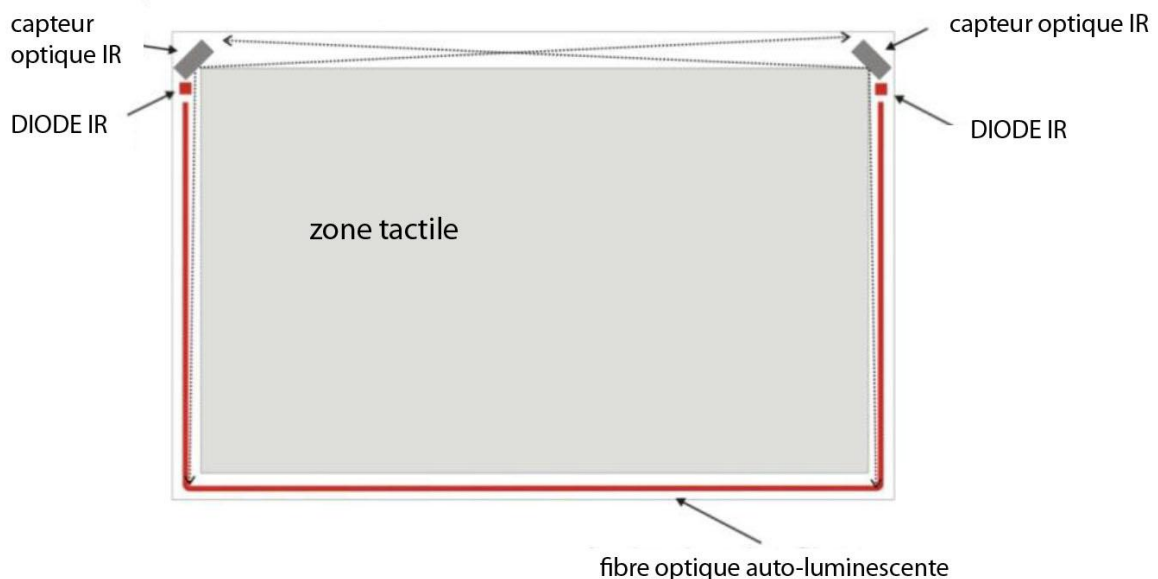


Figure 46 - Illustration simplifiée du fonctionnement d'un cadre optique infrarouge.

3. Périphériques de sortie

Par « réponses système » (*system response*), Maria Karam and M.C. Schraefel sous-entendent de catégoriser le périphérique de sortie du système en fonction du canal utilisé. Il s'agit d'identifier les retours utilisateur et les modalités de sortie du dispositif en situation d'interaction gestuelle. Nous préférons donc utiliser le terme plus général « périphériques de sortie » à la traduction littérale « réponse système¹⁰⁹ ». Parmi les trois catégories avancées par les auteurs¹¹⁰, je retiendrai la sortie visuelle et audio mais j'écarterai celle de la sortie « *commande directe* » qui fait référence aux systèmes de reconnaissance gestuelle permettant d'envoyer des commandes à un autre dispositif ou à une autre application. C'est le cas par exemple des interfaces de contrôle pour des environnements de type ubiquitaire où le geste permet à la fois de désigner un dispositif et d'exécuter ses fonctions à distance. Nous porterons notre attention uniquement aux systèmes dont la sortie est visuelle sachant que cette étude vise à définir un modèle d'interaction et de manipulation graphique par une interface multi-tactile. Toutefois, nous ajouterons une catégorie supplémentaire : la configuration du dispositif, concernant certaines propriétés physiques qui, dans notre cas précis, jouent un rôle important dans les modalités d'interaction entre l'utilisateur et le système.

3.1. Canal visuel

La plupart des systèmes tactiles et gestuels sont équipés d'une sortie visuelle. Afin de préciser davantage mon corpus d'analyse, je limiterai ces dispositifs à ceux dont la manipulation des objets graphiques se fait directement sur l'image et non de manière indirecte comme c'est le cas avec les tablettes graphiques et les souris multi-tactiles. Les écrans se déclinent aujourd'hui en deux grandes familles, les écrans à deux dimensions (les écrans traditionnels) ou trois dimensions : stéréoscopiques (avec lunettes) ou autostéréoscopiques (sans lunettes). J'écarterai également cette dernière famille car d'une part, la plupart des écrans grand format 3D autostéréoscopiques ne fonctionnent

¹⁰⁹ « *System response* ». In KARAM Maria et Schraefel M.C. *A taxonomy of gestures in human computer interaction*. In *Transactions on Computer-Human Interactions*, TBD, 2005

¹¹⁰ KARAM Maria et Schraefel M.C. *A taxonomy of gestures in human computer interaction*. In *Transactions on Computer-Human Interactions*, TBD, 2005

correctement qu'en restant à une certaine distance et sont donc inappropriés pour des interactions tactiles. D'autre part, le couplage entre les représentations tridimensionnelles de l'interface et la surface bidimensionnelle de l'interaction tactile soulèvent des problématiques que nous n'abordons pas ici. Notre modèle d'interaction s'intéresse à l'ensemble des écrans tactiles multipoints, en deux dimensions et de configuration variable.

3.2. Canal audio

La quasi-totalité des dispositifs concernés par notre étude sont équipés d'un système de diffusion audio. Dans certains cas, comme la table musicale *Reactable*, ou le *tenori-on*, il est possible que la diffusion sonore ou musicale constitue le périphérique de sortie principal du dispositif. Cependant, j'écarte de l'étude tout dispositif dont la sortie audio incarnerait l'unique périphérique de sortie. Il pourrait sembler anodin de caractériser les dispositifs tactiles par leur capacité ou non à produire du son car à première vue, il ne sert généralement qu'à accompagner les interactions utilisateurs et bien souvent son utilisation est hasardeuse et précaire. Toutefois, nous verrons plus loin que le canal sonore joue un rôle important dans l'augmentation du feedback haptique.

3.3. Configuration

Selon les besoins et les usages, les dispositifs tactiles se déclinent sous de multiples formes dont les différences affectent directement le type de technologie utilisée mais ces différences ont des conséquences significatives sur les modalités d'interaction tactiles. Dans un premier temps, nous pouvons catégoriser ces dispositifs selon la taille de l'écran, disons de la surface tactile, car comme nous le verrons au chapitre IV - 4.2, cette caractéristique influence directement le type de manipulations possibles (au doigt, à une main ou bi-manuelle) mais également les fonctionnalités potentielles (simple ou multi-utilisateurs). Ensuite, il est possible de caractériser ces dispositifs selon l'inclinaison de l'écran. En effet, nous retrouvons généralement trois configurations possibles : horizontale, verticale ou inclinée. Cette caractéristique a une influence directe sur l'ergonomie, comme la pénibilité de la position verticale ou les zones d'accessibilité de l'écran. Enfin, nous pouvons décliner les dispositifs selon leur portabilité car devoir tenir un écran tout en le manipulant influence considérablement l'ergonomie des interactions.

4. Le canal gestuel

La recherche et le développement d'interfaces post-WIMP répond à la volonté de simplifier les interactions, d'accélérer les processus d'apprentissage et de créer des interfaces plus « intuitives » dans le sens où ces interfaces tentent de réduire le « *fossé d'exécution* » décrit par Hutchin et al.¹¹¹, « *le fossé entre les intentions de l'utilisateur et les actions nécessaires à effectuer sur le dispositif d'entrée* ». Pour y parvenir, les chercheurs se basent sur les capacités perceptuelles et cognitives acquises par l'homme au cours de l'évolution et les moyens offerts par les technologies pour exploiter ces compétences. En diminuant la charge cognitive nécessaire à une tâche spécifique et en donnant à l'utilisateur la possibilité de manipuler les objets virtuels comme il le ferait avec des objets réels, le temps d'apprentissage est également réduit. Les interactions supposent moins de commandes complexes à connaître, comme les raccourcis clavier, et gagnent en immédiateté.

L'émergence de ces interfaces est directement liée aux travaux de recherche qui ont porté une attention toute particulière aux capacités perceptuelles humaines à manipuler des instruments réels ou mixtes. Les interfaces tactiles ayant la particularité de faire disparaître l'instrument physique au profit d'une manipulation manuelle et gestuelle, le « geste à nu » devient le canal privilégié des interactions avec le système. Afin d'apporter une lecture plus précise sur les processus physiques et cognitifs impliqués dans les interfaces qui utilisent le geste dans la manipulation de l'instrument physique ou virtuel, il est nécessaire de revenir sur la richesse du canal gestuel dans les interactions de l'homme avec son environnement.

L'homme a été homo-faber (homme fabricant), avant d'être homo-sapiens (homme savant), explique Henri Bergson¹¹², il a fabriqué des outils parce qu'il n'avait pas suffisamment de qualités physiologiques pour s'adapter convenablement au milieu naturel. A. Leroi-Gourhan, dans « *Le geste et la parole*¹¹³ », en remontant aux origines de l'homme, consacre au geste une étude anthropologique qui nous donne un premier éclairage biologique et sociologique sur le sujet. Il recherche le rapport qui existe entre l'outil et l'artisan ou entre l'outil et l'utilisateur. « *Nous percevons notre pensée comme un bloc et nos outils comme le noble fruit de notre pensée. L'Australanthrope, lui, paraît avoir possédé ses*

¹¹¹ HUTCHINS Edwin L., HOLLAN James D., et NORMAN Donald A. *Direct manipulation interfaces*. In *Human-Computer Interaction* (1, 4), décembre 1985, pp. 311-338

¹¹² BERGSON Henri. *L'évolution créatrice*. PUF. Paris, 2003 (10e édition), 384 p. (coll. Quadrige, n°1)

¹¹³ LEROI-GOURHAN André. *Le geste et la parole*. Paris, Albin Michel, 1964, 2 vol., 323 et 285 p.

outils comme des griffes. » Les termes « outil » et « instrument », en partageant la racine étymologique grecque *organon*, conforte l'hypothèse avancée par Leroi-Gourhan considérant l'outil comme un organe prothétique fonctionnel, comme le prolongement de la main.

Aujourd'hui, nous utilisons des outils appelés machines, qui possèdent leurs propres mécanismes, leurs propres organes et une logique interne. Hannah Arendt précise que cette transition est d'une grande signification car désormais, le corps devient dépendant de ce qu'il utilise, et plus précisément de son rythme autonome¹¹⁴. Leroi-Gourhan démontre à travers son livre, l'extraordinaire évolution du rôle du geste et de la main dans les processus de manipulation, d'expression et de communication. « *La main à l'origine était une pince à tenir les cailloux, le triomphe de l'homme a été d'en faire la servante de plus en plus habile de ses pensées de fabricant. Du Paléolithique supérieur au XIXe siècle, elle a traversé une interminable apogée.*¹¹⁵ »

Par « geste à nu », j'entends souligner l'approche singulière du geste impliqué dans les interfaces multi-tactiles, qui s'affranchit des périphériques mécaniques de saisie pour revenir à des formes plus libres du mouvement. La partie physique de l'instrument d'interaction se déplace vers la main pour profiter de toute la richesse de cet organe et des mouvements qui l'accompagnent. « *La pince à tenir les cailloux* » de l'homme préhistorique a évolué en un outil complexe, dont nous allons, à travers les prochains chapitres, explorer les différentes facettes.

4.1. Définition

Il convient dans un premier temps de distinguer le terme « geste » à celui de « mouvement » en ce sens ou le geste relève d'une intention ou d'un état psychologique. Le dictionnaire le *petit Robert* nous en donne la définition classique suivante : « *mouvement du corps (principalement des bras, des mains, de la tête) volontaire ou involontaire, révélant un état psychologique, ou visant à exprimer, à exécuter qqch.* »

¹¹⁴ ARENDT Hannah. *Condition de l'homme moderne*. Calmann-Lévy. Paris, Pocket, 1983, 406 p.

¹¹⁵ LEROI-GOURHAN André. *Le geste et la parole*. Paris, Albin Michel, 1964, 2 vol., 323 et 285 p.

En revanche , la recherche étymologique¹¹⁶ du terme nous précise que « geste » est un emprunt au latin *gestus* « attitude », « mouvement du corps », « mimique, jeu », et de *gestum*, supin de *gerere* au sens d' « accomplir, faire ». Cette acception se retrouve chez Hugues de Saint-Victor au XIIe siècle qui définit le geste comme « *le mouvement et la figuration des membres du corps adaptés à toute action et attitude.*¹¹⁷ » Cette définition, relativement moderne pour l'époque, renvoie à des notions essentielles pour notre recherche comme l'aspect kinésique et formel (le geste comme mouvement et comme attitude), symbolique (figuration) et pragmatique (le geste lié à l'action). Une distinction entre geste naturel et geste conventionnel s'opère dès la renaissance, notamment grâce au renouveau de la rhétorique antique. Ce n'est qu'au XVIIe siècle, avec la contribution de John Bulwer, « *Chirologia : or the naturall language of hand* », que le geste prend une dimension toute particulière dans les processus de communication. Il y décrit le geste selon une grammaire propre en ayant observé et analysé son apport dans la communication verbale. Il démontre également l'importance de la main dans le geste de communication dont le langage est, selon Bulwer, facilement perceptible et compréhensible, comme si « *l'homme avait une autre bouche ou un moulin à paroles dans sa main.*¹¹⁸ ». La notion de « gestualité » peut alors se comprendre comme l'ensemble des gestes qui sont perçus en tant que signes, actifs ou symboliques.

Le geste est si riche qu'il accompagne non seulement l'activité de communication des hommes mais il traverse également un large champ d'activités liées à l'expression : l'écriture, la peinture, la sculpture, le théâtre, la musique. Quintilien rapprochait par exemple le geste de rhétorique à celui de l'art pictural en décrivant le pouvoir expressif du geste artistique :

« Et il n'est pas étonnant que ces gestes, qui, après tout, comportent quelque mouvement, produisent une impression si profonde sur l'âme, quand la peinture, qui est un ouvrage silencieux et qui fixe des attitudes immuables, agit sur notre

¹¹⁶ Dictionnaire Historique de la Langue Française. Dictionnaire historique. Sous la direction d'Alain Rey. Le Robert, 2006, 4304 p.

¹¹⁷ « *Gestus est motus et figuratio membrorum corporis, ad omenem agendi et habendi modum* ». In SCHMITT Jean-Claude. *La raison des gestes dans l'Occident médiéval*. Paris, Gallimard, 1990, 432 p.

¹¹⁸ « [...] man had another mouth or fountain of discourse in his hand ». In BULWER John. *Chirologia or The Natural Language of the Hand*. Kessinger Publishing Co, 2003, 380 p.

*sensibilité la plus intime au point d'avoir parfois l'air d'être plus éloquente que la parole même.*¹¹⁹ »

Jean-Claude Shmitt, dans un ouvrage intitulé « *La Raison des Gestes*¹²⁰ » reprend dans sa typologie du geste les fonctions de communication mais étend le geste de l'action à la dimension du « faire », répondant à des nécessités matérielles ou artistiques et à sa dimension expressive, comme expression des mouvements intérieurs de l'âme, des sentiments et de la vie morale de l'individu.

La polysémie du terme nous conduira dans la suite de ce chapitre à analyser les taxinomies du geste issues de la psychologie, de la linguistique et des sciences informatiques afin d'obtenir un cadre de compréhension général du geste et une lecture plus fine des caractères du geste impliqués dans les interfaces utilisateur multi-tactiles.

4.2. Intégration sensori-motrice

L'intégration sensori-motrice se définit comme les processus qui organisent les sensations du corps à son environnement. Plus spécifiquement, c'est la capacité du cerveau à traiter plusieurs modalités sensibles en entrée et de les convertir en sorties fonctionnelles. Pendant longtemps, les scientifiques ont pensé qu'à chaque organe sensible correspondait une zone spécifique de traitement dans le cerveau. Aujourd'hui, nous savons¹²¹ que ces différentes zones ne sont pas forcément spécialisées mais qu'elles peuvent utiliser plusieurs modalités de perception afin d'analyser les informations issues de nos sens et permettre au corps d'évoluer et de comprendre notre environnement. L'intégration sensori-motrice fait appel aux cinq sens connus que sont la vue, l'ouïe, le toucher, l'odorat et le goût. D'autres modalités sensorielles sont également impliquées dans les processus de perception comme le système vestibulaire (sens de l'équilibre) et la proprioception (sens de la position d'un organe dans l'espace).

¹¹⁹ « *Nec mirum, si ista, quae tamen in aliquo posita sunt motu, tantum in animis valent, cum pictura, tacens opus et habitus semper eiusdem, sic in intimos penetret adfectus, ut ipsam vim dicendi nonnunquam superare vidatur.* » QUINTILIEN. *Institution oratoire*. Traduction : Jean Cousin. Paris. Belles Lettres. 1979, (Coll. Universités de France)

¹²⁰ SCHMITT Jean-Claude. *La raison des gestes dans l'Occident médiéval*. Paris, Gallimard, 1990, 432 p.

¹²¹ MACALUSO Emiliano et DRIVER Jon. *Multisensory spatial interactions : a window onto functional integration in the human brain*. In *Trends in Neurosciences* (28): 2005, pp. 263-271

Cette courte introduction me permet d'ouvrir une petite fenêtre sur l'histoire et de raconter l'origine des caricatures humaines de la perception connues sous le nom d'homoncules sensoriels et moteurs. Dans les années trente, le docteur Wilder Penfield, pionnier de l'intégration des principes neurophysiologiques dans la pratique neurochirurgicale, cherchait une solution aux problèmes des convulsions causées par l'épilepsie. Il utilisa une électrode afin de stimuler différentes régions du cortex cérébral et demanda à ses patients, dans un état conscient, ce qu'ils ressentaient. La « carte » anatomique qu'il réalisa demeure la première représentation de l'intensité sensorielle dérivée des différentes parties du corps humain. Cette carte a donné lieu à deux représentations déformées du corps humain (cf. figure 47 et 48) dont la vue est tout à fait troublante, où les parties du corps sont mises en relation avec la taille des régions corticales correspondantes : l'homoncule moteur qui désigne l'organisation de la commande motrice des muscles du corps humain et l'homoncule sensitif, ou somesthésique qui désigne la finesse, la complexité et la richesse de la sensibilité des parties du corps.

Ce que je souhaite mettre en exergue à travers ces représentations est la zone de la main dont la surface sur le territoire corticale est étonnamment surdimensionnée par rapport aux autres membres du corps. Ce premier indice nous renseigne sur la complexité et à la richesse non seulement motrice mais sensitive de cet organe. Ce sont les premiers indices qui nous permettent de prendre la mesure de l'implication de la main dans notre activité perceptivo-motrice. Les facultés que nous avons de percevoir notre environnement, de transformer la matière, de ressentir la texture des objets et de communiquer reposent en grande partie sur la richesse des interactions avec le monde dont notre main est capable.

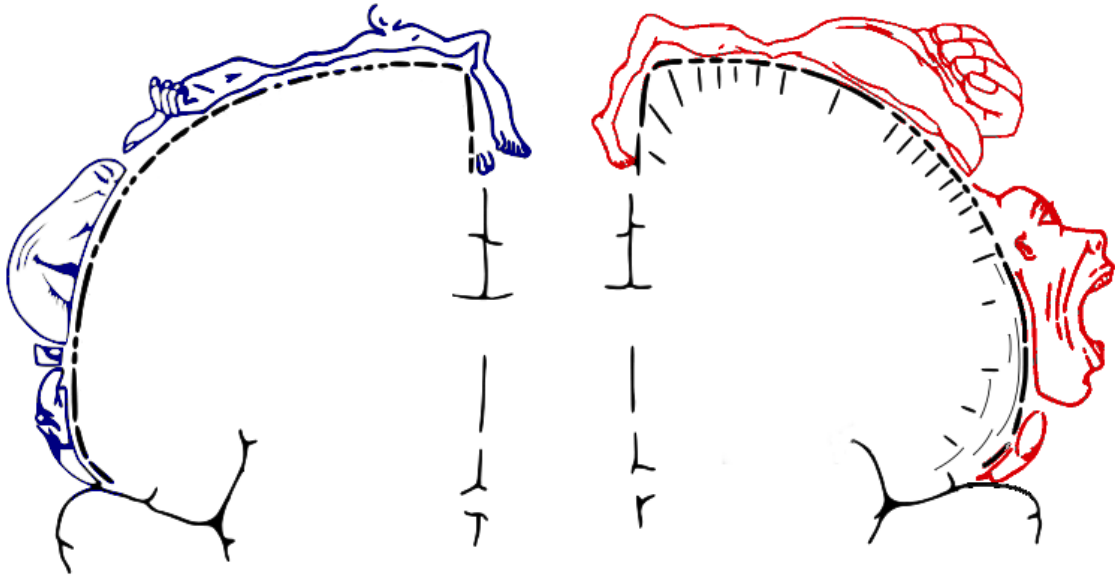


Figure 47 - Homoncule sensitif montrant les différentes régions du cortex cérébral correspondant à l'intensité et à la richesse des terminaisons nerveuses des différentes parties du corps.



Figure 48 - Sculpture de l'homoncule moteur exposée au muséum d'Histoire Naturelle de Londres. L'homoncule sensitif est quasiment identique à celui-ci.

4.3. Continuum de conventionnalité

John Bulwer, physicien anglais, propose, dès 1644, une classification¹²² des mouvements de la main accompagnant le discours. Il publiait déjà une véritable grammaire gestuelle consacrée à la rhétorique du geste et de la voix. Bulwer fait la description de quarante-neuf gestes de la main et trente positions des doigts définis comme les canons de l'art oratoire. Plus récemment, Nespoulous et Lecours¹²³ divisent les gestes selon leur intentionnalité communicationnelle. Les gestes centrifuges sont directement adressés à un interlocuteur tandis que les gestes centripètes correspondent aux gestes observables mais ne s'adressent à personne en particulier. Or, dans le cas des IHM multi-tactiles, nous partons du postulat que tout geste effectué sur la surface tactile possède une intention communicationnelle destinée à la commande du système ou à la manipulation d'objets. Contrairement au système *VideoDesk* par exemple, nous n'avons pas la contrainte, soulevée par Jakob Nielsen¹²⁴, de devoir distinguer les gestes intentionnels des gestes de gesticulation :

« Un des problèmes qu'ils avaient en développant leur langage gestuel était de déterminer si les mouvements de votre main étaient parce que vous vouliez la déplacer de l'autre côté de l'écran ou s'il s'agissait d'une commande. »¹²⁵

En considérant la distinction proposée par Nespoulous et Lecours, nous écartons donc les gestes inconscients ou ceux révélant d'un état psychologique en orientant notre recherche vers les gestes centrifuges, intentionnels.

McNeill démontre en 1992 que le geste et le langage ne sont pas équivalents mais qu'ils sont complémentaires. Considérer le geste comme celui qui accompagne la parole (communication co-verbale) ou qui s'y substitue (langue des signes) nous incite à distinguer

¹²² BULWER John. *Chirologia or The Natural Language of the Hand*. Kessinger Publishing Co, 2003, 380 p.

¹²³ NESPOULOUS J-L et LECOURS A. R. *Gestures: Nature and Function*. In *The Biological foundations of Gestures : Motor and Semiotic Aspects..*). Lawrence Erlbaum Associates, Hillsdale, 1986, pp. 49-62

¹²⁴ NIELSEN Jakob. *CHI'88 Trip Report*. Washington D.C., 15-19 mai 1988

¹²⁵ « One problem they had in developing their gestural language was in parsing hand movements to determine when you just want to move your hand to another part of the screen and when you want to issue a command. » In NIELSEN Jakob. *CHI'88 Trip Report*. Washington D.C., 15-19 mai 1988. Traduction libre COQUELIN Marine et MORIN Rudy.

les différents degrés de conventionnalité impliqués dans la communication gestuelle. Le « Continuum de Kendon¹²⁶ » propose la classification suivante :

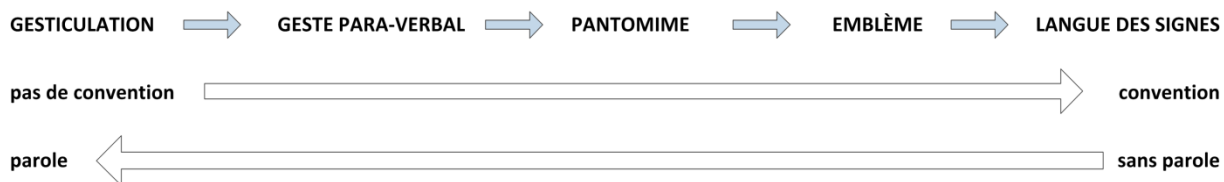


Figure 49 - Représentation du continuum de Kendon. Reproduction. In [McNEILL, 1996].

La gesticulation. Il s’agit de mouvements spontanés de la main ou du bras. Ces gestes sont également appelés *illustrateurs* par Paul Ekman¹²⁷ et Wallace Friesen. Ce sont des mouvements idiosyncrasiques, propres au comportement de l’individu, qui accompagnent toujours la parole et qui sont produits spontanément. Ces gestes sont généralement réalisés de manière naturelle et inconsciente. Leur degré de conventionnalité est quasi nul.

Le geste para-verbal (*language-like gestures*). Ils dépendent d’un système de codage au vocabulaire limité. Il s’agit par exemple des gestes remplaçant certains mots dans une phrase. Cette catégorie de geste étant peu pertinente, elle a été supprimée par la suite lorsque d’autres chercheurs ont fait référence au continuum de Kendon.

Pantomime. Il s’agit d’un geste ou d’une séquence de gestes qui ont une fonction narrative. Ils ne sont pas obligatoirement accompagnés par la parole et ils présentent des fortes propriétés iconiques afin d’être facilement identifiés par l’individu à qui ils sont destinés.

Emblèmes. Ce sont des gestes qui présentent un fort degré de conventionnalité car ils remplacent souvent la parole et correspondent à des codes culturels. C’est par exemple le geste de salut ou celui qui signifie l’affirmation ou la négation.

Langue des signes. Il s’agit de la langue utilisée par les personnes atteintes de surdité afin de communiquer. Ce type de geste assure toutes les fonctions remplies par la parole et dépend d’un code gestuel très précis. Ainsi le degré de conventionnalité de ce geste est important afin d’assurer une compréhension optimale entre les individus.

¹²⁶ McNEILL David. *Hand and Mind: What Gestures Reveal about Thought*. University Of Chicago Press, 1996, 423 p.

¹²⁷ EKMAN Paul et FRIESEN Wallace V. *The repertoire of nonverbal behavior: Categories, origins, usage, and coding*. In *Semiotica* (janvier 1969), pp. 49- 98

4.4. Taxinomie fonctionnelle du geste

Le geste a une particularité qui le distingue des autres canaux de communication, c'est qu'il possède à la fois les moyens d'action et de transformation du monde physique mais il est également un moyen de communication en émettant et en recevant des informations. Le geste est en effet pourvu du sens du toucher qui implique des systèmes perceptifs et moteurs.

« La main est une structure complexe dont la fonction perceptive est imbriquée à une fonction motrice (...): celle d'instrument de travail grâce auquel sont réalisés la saisie, le transport et la transformation des objets à des fins utilitaires.¹²⁸ »

Les méthodes de perception distinguent différents types de gestes, catégorisés par Claude Cadoz en trois classes fonctionnelles¹²⁹. D'une part la fonction épistémique, c'est-à-dire le geste du toucher, celui qui vise à connaître ou à reconnaître. Il s'agit de la conjonction de perceptions tactilo-kinesthésiques (le toucher, la pression, la traction) et de perceptions proprioceptives (muscles et articulations). D'autre part, la fonction sémiotique est celle qui sert à faire connaître, à désigner, à communiquer. Le geste qui accompagne la parole n'est pas uniquement illustratif, il en dit parfois plus sur les émotions que la voix elle-même. Puis il existe la fonction ergotique (de « *ergon* », travail en grec), celle qui désigne les gestes dont le rapport avec l'environnement n'est pas uniquement informationnel mais aussi énergétique.

Selon Quek *et al.*¹³⁰, pour que les interactions homme-machine se rapprochent du niveau de transparence des interactions entre les hommes, il est nécessaire de comprendre la phénoménologie des interactions de communication et les classes de caractères que l'on peut extraire des gestes pour en avoir une meilleure compréhension. Afin d'apporter un regard plus précis sur le geste en situation d'interaction tactile et gestuelle, nous allons préciser quel est le degré d'implication des gestes sémiotiques dans les interactions tactiles

¹²⁸ HATWELL Yvette. *Toucher l'espace. La main et la perception tactile de l'espace*. Lille, Presses Universitaires de Lille. 1986, 374 p.

¹²⁹ CADOZ Claude. *Le Geste canal de communication homme-machine : la communication instrumentale*. In *TSI (Traitement du Signal et des Images)*, vol. 13. 1994, pp. 31-61 (coll. Hermès)

¹³⁰ QUEK Francis, McNEILL David, BRYLL Robert, DUNCAN Susan, MA Xin-Feng, KIRBAS Cemil, McCULLOUGH Karl E. et ANSARI Rashid. *Multimodal human discourse: gesture and speech*. ACM, Computer-Human Interact (9, 3), 2002, pp. 171-193

multi-points. Nous verrons à travers une étude non exhaustive des contributions existantes de quelle manière le geste est engagé dans les modèles d'interactions et quelles en sont les composantes principales. Les théories sur le geste couvrent une multitude de champs disciplinaire comme l'anthropologie, la linguistique, les sciences cognitives et la psychologie.

Cette étude sur le geste nous donnera un cadre de compréhension des modalités d'interactions manuelles qui guidera notre réflexion sur la définition du modèle d'interaction multi-tactile et le développement de notre environnement de développement. Est-il pertinent de connaître la pression des doigts, l'orientation de la main, l'angle et la vélocité des mouvements ? Quels sont les gestes spécifiquement impliqués dans l'interaction tactile ? Existe-t-il des modalités d'interaction plus simples ou plus naturelles ? En détaillant les approches théoriques issues de la sémiologie et de la linguistique, nous obtiendrons une classification générale du geste en situation de communication entre les hommes. Comme le font remarquer Maria Karam et M.C. Schraefel¹³¹, les termes désignant le même type de geste sont différents d'une taxinomie à une autre. Par exemple, la gesticulation décrit des gestes également désignés comme co-verbaux¹³², pantomimes¹³³ ou naturels¹³⁴. Le geste naturel est parfois appelé geste libre¹³⁵ et le geste symbolique est parfois nommé iconique ou geste tracé en fonction des auteurs. Nous tenterons d'établir une taxinomie qui permette de rendre compte des principales contributions existantes et d'en éclaircir les termes.

4.4.1. Fonction épistémique

Le sens du toucher est réparti sur l'ensemble du corps. Les informations du toucher transmises au cerveau concernent trois types de sensation : la température, le contact et la douleur. Cependant, le sens du toucher fait partie d'un système sensoriel de l'organisme plus large appelé la somesthésie. Il s'agit d'un ensemble de sensations du toucher qui provient d'un nombre important de capteurs situés dans différentes régions du corps. Ces

¹³¹ KARAM Maria et Schraefel M.C. *A taxonomy of gestures in human computer interaction*. In *Transactions on Computer-Human Interactions*, TBD, 2005

¹³² KETTEBEKOV Sanshazar. *Exploiting prosodic structuring of coverbal gesticulation*. In *Proceedings of the 6th international conference on Multimodal interfaces (ICMI '04)*. ACM, New York, USA, 2004, pp. 105-112

¹³³ QUEK Francis, McNEILL David, BRYLL Robert, DUNCAN Susan, MA Xin-Feng, KIRBAS Cemil, McCULLOUGH Karl E. et ANSARI Rashid. *Multimodal human discourse: gesture and speech*. ACM, Computer-Human Interact (9, 3), 2002, pp. 171-193

¹³⁴ WEXELBLAT Alan. *Natural gesture in virtual environments*. In *Proceedings of the conference on Virtual reality software and technology (VRST '94)*. World Scientific Publishing Co., USA, 1994, pp. 5-16

¹³⁵ HARDENBERG Christian von et BÉRARD François. *Bare-hand human-computer interaction*. In *Proceedings of the 2001 workshop on Perceptive user interfaces (PUI '01)*. ACM, New York, USA, 2001, pp. 1-8

sensations proviennent des tissus de l'organisme comme le derme, les viscères, les muscles et les tendons. La combinaison de ces informations est à l'origine de sensations corporelles étroitement liée à la perception de l'espace et des formes. La stimulation de la peau associée à la palpation et aux mouvements exploratoires de la main renseignent le cerveau sur la forme, l'orientation, la distance, la grandeur d'un objet, etc. grâce aux multiples informations envoyées par les cellules nerveuses des organes engagés dans le geste. Ces processus complexes forment un ensemble indissociable appelé perception tactilo-kinesthésique ou haptique. Bien que la main soit l'organe qui possède le plus de terminaisons nerveuses, ce qui en fait l'organe privilégié pour le sens tactile, c'est la combinaison de l'ensemble des fonctions spécialisées dans le traitement des sensations corporelles qui nous permet d'appréhender l'objet à percevoir dans son intégralité. Alors que le sens du toucher implique la réception passive¹³⁶ d'informations cutanées comme la pression, la température et la douleur, le sens kinesthésique correspond à la perception des mouvements alors que le sens proprioceptif renseigne sur la position des articulations et du corps.

La vision et l'haptique diffèrent, selon Revesz¹³⁷, en ce sens où à travers la vision, l'impression de l'objet est instantanée alors qu'à travers la perception tactilo-proprio-kinesthésique, la représentation de l'objet se fait (à différents degrés) par l'assemblage de sensations tactiles discrètes. Les objets dans le monde physique sont multidimensionnels et leurs propriétés s'inscrivent dans toutes ces dimensions. Le mode exploratoire de la perception haptique implique une perception séquentielle de la taille, des textures et des formes de l'objet, il s'agit d'une perception analytique qui s'opère dans le temps de l'exploration. Ce type de perception engagerait davantage de processus cognitifs que l'exercice de la vision. Notre connaissance du monde est le résultat de l'expérience de notre environnement par les sens. La perception haptique y participe tout autant que la vision si ce n'est davantage car elle peut quasiment s'y substituer et fournir des informations que la vue seule ne saurait renseigner.

« L'importance épistémologique de l'haptique devient évidente quand nous réalisons que nous sommes confrontés à un espace qui est tout autant

¹³⁶ GIBSON J. James. *Observations on active touch*. In *Psychological Review* (69), 1962, pp. 477–490

¹³⁷ RÉVÉSZ Géza. *Psychology and Art of the Blind*. London, Longmans Green, 1950, 338 p.

*autonome que l'espace visuel, et que nos sensations spatiales haptiques sont aussi originales que nos sensations spatiales.*¹³⁸ »

Dans « *La Science et l'Hypothèse* », Poincaré nous invitait déjà à saisir le concept d'espace par une approche tactile et musculaire du geste.

*« Chaque muscle donne naissance à une sensation spéciale susceptible d'augmenter ou de diminuer, de sorte que l'ensemble de nos sensations musculaires dépendra d'autant de variables que nous avons de muscles. De ce point de vue, l'espace moteur aurait autant de dimensions que nous avons de muscles.*¹³⁹ ».

Le concept d'espace renvoie donc inévitablement au geste, au geste intériorisé. Les sensations liées à nos mouvements exploratoires constituent les signaux somesthésiques des fonctions tactiles, kinesthésiques et proprioceptives qui contribuent à la genèse de cet espace.

4.4.2. Fonction ergotique

La fonction ergotique du geste invite à considérer la main comme un organe d'action doté à la fois d'une fonction motrice et perceptive. La main est un organe qui engage le corps dans l'action, le transport et la transformation de la matière. Dans ce cas, la main ne communique pas d'information mais de l'énergie et des forces qui sont appliquées à l'objet en le déplaçant ou en le transformant, induisant un retour d'énergie passif (densité de l'objet) ou actif (retour de force ou mouvement intrinsèque). Ce qui caractérise l'action motrice de la main selon Claude Cadoz¹⁴⁰, c'est « *qu'elle est en prise directe avec la matière, qu'elle peut la modeler, la transporter, l'usiner, la briser.* » La fonction ergotique de la main nous rappelle que la main est avant tout un outil, un instrument (*organon* en grec), non spécialisé et

¹³⁸ « *The epistemological importance of Haptics becomes clear when we realise that we are here dealing with a space which is just as autonomous as the visual space, and that our haptic spatial impressions are possessed of the same originality as the visual spatial impressions.* » In *ibid.* Traduction libre : COQUELIN Marine et MORIN Rudy.

¹³⁹ POINCARÉ Henri. *La science et l'hypothèse*. Flammarion. Paris, 1968, 252 p. (coll. Champs Flammarion)

¹⁴⁰ CADOZ Claude. *Le Geste canal de communication homme-machine : la communication instrumentale*. In *TSI* (Traitement du Signal et des Images), vol. 13. 1994, pp. 31-61 (coll. Hermès)

capable de répondre à des usages de manipulation ou de maniement. « *La main semble bien être non pas un outil, mais plusieurs. [...] C'est donc l'être capable d'acquérir le plus grand nombre de techniques que la nature a donné l'outil de loin le plus utile, la main.*¹⁴¹ » Cependant, le développement récent des théories énavivistes nous amènent à supposer que la fonction ergotique de la main ne peut être dissociée de sa fonction épistémique en ce sens où ce que nous percevons est le résultat de l'interdépendance entre les actions que nous réalisons et les sensations qui en résultent. Noë et Lenay¹⁴² avancent l'hypothèse que la perception, en tant qu'expérience du réel, ne répond pas à une réception passive mais qu'elle se construit et s'actualise en fonction de la contingence sensorimotrice, c'est à dire la loi qui unit l'action à ses effets.

Le geste ergotique ou manipulatif (Quek *et al.*) peut être classifié selon des caractéristiques physiques comme le type d'objet manipulé, le type de modification apportée à l'objet, combien de personnes ou de mains sont impliquées et le niveau d'indirection de la manipulation (utilisation d'un outil par exemple). Il est également possible de classer ces gestes selon leur fonction et les modes de préhension qu'ils induisent. Quel *et al.* définissent le geste ergotique pour le contrôle d'un objet comme « *une relation étroite entre les mouvements du geste (main/bras) avec l'entité manipulée.*¹⁴³ » La préhension est la faculté ou l'action de saisir un objet, soit pour en apprécier ses caractéristiques (dimension, poids, forme), soit pour le transporter ou le manipuler. McKenzie¹⁴⁴ définit la préhension comme « *l'application de forces fonctionnelles efficaces par la main à un objet pour une tâche, en fonction d'un certain nombre de contraintes.* » La classification proposée par McKenzie *et al.*¹⁴⁵ rend compte de la relation entre d'une part les propriétés musculaires et articulaires de la main et d'autre part les caractéristiques de la tâche, comme la précision et la force qui sont nécessaires à sa réalisation.

¹⁴¹ ARISTOTE. *Les Parties des animaux*. Belles Lettres, 2003, 193 p. (Coll. Universités de France).

¹⁴² O'REGAN J. Kevin et NOË Alva. *A sensorimotor account of vision and visual consciousness*. In *Behavioral and brain sciences*, 2001, pp. 939-1031

¹⁴³ QUEK Francis, McNEILL David, BRYLL Robert, DUNCAN Susan, MA Xin-Feng, KIRBAS Cemil, McCULLOUGH Karl E. et ANSARI Rashid. *Multimodal human discourse: gesture and speech*. ACM, Computer-Human Interact (9, 3), 2002, pp. 171-193

¹⁴⁴ « [...] *the application of functionally effective forces by the hand to an object for a task, given numerous constraints* » In MCKENZIE C. L. et IBERALL T. *The Grasping hand*. In *Advances in Psychology* (104). 1994, 500 p. Traduction libre : COQUELIN Marine et MORIN Rudy.

¹⁴⁵ MCKENZIE C. L. et IBERALL T. *The Grasping hand*. In *Advances in Psychology* (104). 1994, 500 p.

Par exemple, la préhension est une notion largement impliquée dans les interfaces utilisateur tangibles car les instruments d'interactions peuvent être attrapés, déplacés et manipulés. C'est également le cas dans les interfaces de réalité virtuelle où la préhension d'objets graphiques est une action fréquemment réalisée, grâce à des gants de données par exemple, mais à un moindre degré dans le sens où la manipulation des objets est virtuelle et n'atteint pas le même niveau de réalisme. Enfin, les interfaces utilisateur tactiles sont un cas particulier car elles incitent à la préhension en ce sens où les objets possèdent une affordance perceptible. Toutefois, ces objets demeurent en deux dimensions et ils ne peuvent être saisis comme le seraient des objets réels. Cette dichotomie peut être perturbante pour l'utilisateur qui doit manipuler des objets graphiques sans pouvoir réellement les saisir. C'est un point que nous développerons lorsque nous aborderons les notions d'affordance physique et d'affordance perçue.

4.4.3. Fonction sémiotique

La conversation humaine est riche en interactions verbales et tout particulièrement en interactions non-verbales où les gestes jouent un rôle essentiel car ils accompagnent le discours et apportent de l'information de communication qui permet d'exprimer une idée, de communiquer ou d'accompagner la parole. Selon Francis Quek *et al.*¹⁴⁶, les gestes employés dans la conversation se distinguent des gestes ergotiques de plusieurs manières.

« Premièrement, car les intentions de ces derniers ont une fin de manipulation, il n'y a aucune garantie que les traits (figures) essentiels de la main soient visibles. Deuxièmement, la dynamique des mouvements de la main en situation de manipulation diffère significativement des gestes conversationnels. Troisièmement, les gestes manipulatifs peuvent être accompagnés par un feedback visuel, tactile ou à retour de force par l'objet en train d'être manipulé,

¹⁴⁶ QUEK Francis K. H. *A purpose taxonomy of gestural interaction*. Vision interfaces and systems laboratory, center for human compute interaction, non publié, 2003. Disponible ici : <http://vislab.cs.vt.edu/~quek/Classes/Aware+EmbodiedInteraction/PAPERS/Que03.pdf>

alors que les gestes sémiotiques sont généralement effectués sans ce type de contraintes.¹⁴⁷ »

Les gestes sémiotiques peuvent accompagner la parole ou au contraire se suffire comme unique moyen d'expression. Dans le premier cas, nous parlerons de geste co-verbal (pantomime ou gesticulation selon les auteurs) alors que le second fait référence à la langue des signes ou aux gestes que l'on utilise dans une situation où la parole n'est pas possible. C'est le cas des signes, hautement conventionnels, que l'on peut produire en plongée ou dans un milieu bruyant afin de communiquer. Cette classe de gestes fait l'objet d'études approfondies en IHM pour le développement d'interfaces gestuelles à distance ou pour concevoir des systèmes capables d'analyser la langue des signes. C'est également le cas des interfaces dites multimodales qui permettent à l'utilisateur d'utiliser plusieurs canaux de communication simultanément comme le geste, la parole, les expressions faciales, le mouvement des yeux, etc. Dans ces conditions, le geste a une définition plus large et concerne l'ensemble des manifestations physiques de l'utilisateur.

De manière générale, Claude Cadoz définit les gestes sémiotiques comme ceux qui ont pour fonction de faire connaître, de produire un message informationnel à destination de l'environnement. Quek *et al.* parlent de gestes « sémaphoriques », en désignant la fonction de communication du geste par sa fonction signalétique. Du grec ancien *sêma* (signe) et de la racine *-phore* (porteur), le geste est alors vecteur d'un signal. « *De tels sémaphores concernent les gestes en pose statique ou les mouvements au style prédéfini.*¹⁴⁸ »

Pour nous faire une idée plus précise des différentes formes du geste sémiotique, nous pouvons les classer selon le niveau d'abstraction du mouvement par rapport à l'intention d'interprétation initiale. Quek *et al.*¹⁴⁹ distinguent deux catégories principales : les gestes symboliques et les gestes actés.

¹⁴⁷ QUEK Francis, McNEILL David, BRYLL Robert, DUNCAN Susan, MA Xin-Feng, KIRBAS Cemil, McCULLOUGH Karl E. et ANSARI Rashid. *Multimodal human discourse: gesture and speech*. ACM, Computer-Human Interact (9, 3), 2002, pp. 171-193

¹⁴⁸ QUEK Francis K. H. *A purpose taxonomy of gestural interaction*. Vision interfaces and systems laboratory, center for human compute interaction, non publié, 2003. Disponible ici :

<http://vislab.cs.vt.edu/~quek/Classes/Aware+EmbodiedInteraction/PAPERS/Que03.pdf>

¹⁴⁹ QUEK Francis K. H. Virtual reality software & technology: proceedings of the VRST '94

Les gestes symboliques ou iconiques (McNeill) concernent les mouvements de la main dont le rôle est principalement linguistique. Le langage des signes fait partie par exemple de cette classe gestuelle. Quek *et al.* insistent sur l'aspect arbitraire de ce type de geste. Il existe en effet une dichotomie au sein des sémiologues sur l'opacité des gestes symboliques et l'universalité de leur interprétation. On ne peut en effet mettre de côté l'influence des origines culturelles de l'homme dans sa façon de communiquer avec les gestes. Même en considérant que certains gestes symboliques sont communs à plusieurs cultures à un certain moment de l'histoire, Nespoulos et Lecours démontrent que leur évolution diachronique divise ses gestes en des fragments opaques.

Quek *et al.* classifient les gestes symboliques en gestes référentiels et modélisants :

- *Les gestes référentiels* permettent de désigner un objet ou un concept. Frotter le doigt et l'index font référence à l'argent ou à la monnaie par exemple. Ils sont nommés *kinétographique* lorsqu'ils font référence à une action, *idéographique* lorsqu'ils renvoient à une idée (kendon) et *pictographiques* (Ekman et Friser¹⁵⁰) lorsqu'ils renvoient à un objet.
- *Les gestes modélisants* sont généralement utilisés en complément du discours pour modéliser le référent ou donner une opinion. Balancer la main de gauche à droite avec l'index levé exprime par exemple le désaccord de la personne. Effectuer une rotation des mains l'index levé peut signifier « continue de donner des exemples » ou « continue d'avancer » et dépend du contexte dans lequel il est réalisé.

Les gestes actifs sont divisés en deux catégories :

- les *gestes mimétiques* sont caractérisés par leur iconicité. Ce sont généralement des gestes effectués afin de mimer le référent. Ce type de geste, appelés également pantomime (Efron), ne répondent généralement à aucune convention particulière et sont réalisés « à la volée ». Par exemple, une personne faisant mine d'allumer une cigarette avec un briquet imaginaire signifie qu'il cherche du feu.

¹⁵⁰ EKMAN Paul et FRIESEN Wallace V. *The repertoire of nonverbal behavior: Categories, origins, usage, and coding*. In *Semiotica* (janvier 1969), pp. 49- 98

- *Les gestes déictiques* sont une classe gestuelle un peu particulière car ils se situent à l'intersection des gestes ergotiques et des gestes sémiotiques. Ils ont d'une part la capacité d'établir la position spatiale d'un objet ou d'en définir la trajectoire en le déplaçant, et d'autre part, ils facilitent la concrétisation d'entités abstraites ou distantes lors de la conversation. Quek *et al.*¹⁵¹ distinguent trois classes différentes, au sein de cette catégorie, déterminées par leur contexte. Les *gestes déictiques de spécification* interviennent lorsque le sujet sélectionne un objet ou une position particulière. Toucher le bouton d'une interface graphique avec le doigt est par exemple un geste de spécification. Les *gestes déictiques génériques* permettent de cibler une classe d'objets en sélectionnant un de ces membres. Enfin, les *gestes déictiques métonymiques* sont réalisés lorsqu'un sujet pointe un objet pour signaler sa relation avec une autre entité ou pour désigner sa fonction.

Les gestes de battement sont définis par Paul Ekman et Frisen¹⁵² comme ceux qui accompagnent le discours par un rythme de la main ou accentuent une syllabe par un mouvement plus important. Ce type de geste n'apporte guère d'informations supplémentaires au dialogue et appartient à la catégorie de gestes définis par Wexelblat comme privés. Opposés aux gestes publics qui ajoutent du sens au discours et sont capables parfois de remplacer la parole, ce type de gestes n'apporte aucun intérêt évident pour le dialogue homme-machine et sera donc ignoré dans la suite de cette étude.

¹⁵¹ QUEK Francis K. H. *Toward a vision-based hand gesture interface*. In *Proceedings of the conference on Virtual reality software and technology (VRST '94)*. World Scientific Publishing Co., USA, 1994, pp. 17-31

¹⁵² EKMAN Paul et FRIESEN Wallace V. *The repertoire of nonverbal behavior: Categories, origins, usage, and coding*. In *Semiotica* (janvier 1969), pp. 49- 98

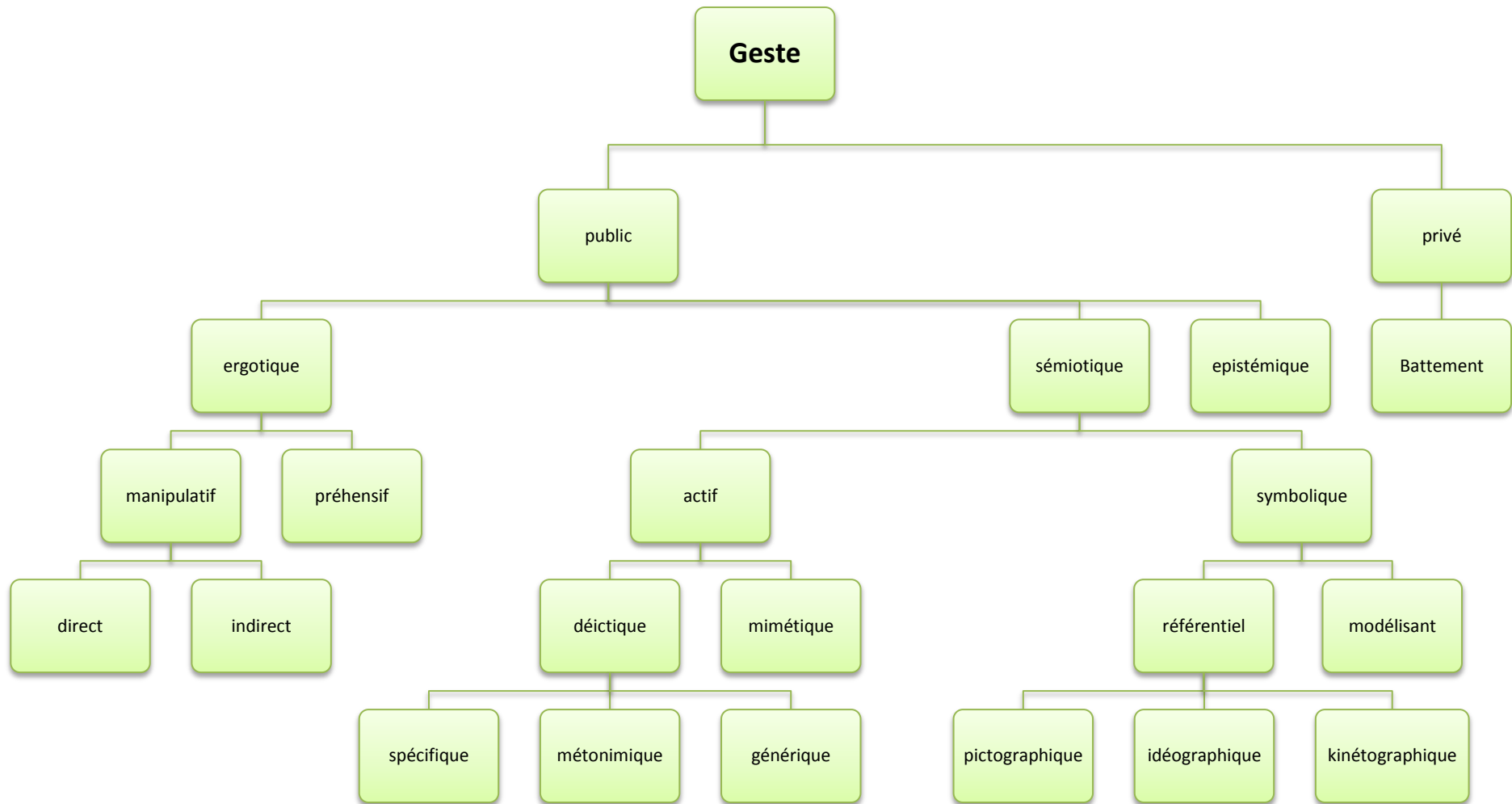


Figure 50 - Taxinomie générale du geste.

CHAPITRE IV MODÈLE D'INTERACTION

« Bien souvent, les exigences des interfaces spécifient que le produit final soit intuitif, ou naturel. Cependant, il n'existe pas de faculté humaine telle que l'intuition, au sens ordinaire du terme ; à savoir, le savoir acquis sans exposition préalable au concept, sans être passé par un processus d'apprentissage, et sans avoir à réfléchir de manière rationnelle. »

« Many interface requirements specify that the resulting product be intuitive, or natural. However, there is no human faculty of intuition, as the word is ordinarily meant; that is, knowledge acquired without prior exposure to the concept, without having to go through a learning process, and without having to use rational thought. »

Jef Raskin

Au cours des années 80, Les contributions scientifiques de Lloyd H. Nakatani et John A. Rohrlich, des laboratoires *Bell*, ont été déterminantes car elles ont tracé la voie pour les recherches et les développements à venir sur les interfaces physiques à base de capteurs. Dans l'article « *Soft Machines : A Philosophy of User-Computer Interface Design*¹⁵³ », les auteurs distinguent les machines des systèmes d'exploitation selon des caractéristiques qui ont une influence considérable sur les modalités d'interaction entre l'utilisateur et les objets d'intérêt de la manipulation. D'un côté, les machines sont dédiées à une tâche spécifique, leur forme suggère fortement leur fonction et leur contrôle est directement lié aux actions réalisées. Au contraire, les systèmes d'exploitation ne sont pas limités à une tâche spécifique, leur forme n'est pas déterminante et l'action sur un clavier ou une souris n'a pas forcément de conséquence immédiate ou prédéterminée. Là où le *hardware* est inflexible et limité à ses prédispositions physiques et fonctionnelles, le *software* « *permet de capitaliser les avantages de ces machines et de dépasser leurs limitations.* »

A partir de ce constat, Nakatani et Rohrlich développent un modèle conceptuel : les *soft-machines* qu'ils définissent comme des interfaces « *utilisant la combinaison en synergie de graphiques numériques temps réels pour l'affichage de contrôles logiciels, et un écran tactile pour rendre ces « contrôles-souples*¹⁵⁴ *» manipulables comme des contrôles physiques conventionnels.*¹⁵⁵ » C'est ce que Bill Buxton entend lorsqu'il évoque l'alternative proposée par les écrans tactiles : « *ce qui peut faire le moins peut faire le plus.*¹⁵⁶ » La physiologie des technologies tactiles, comme interfaces non spécialisées, autorisent de multiples opérations de contrôle tout en réduisant leur apparente complexité. Les écrans tactiles sont comme des interfaces « caméléons » qui peuvent se transformer en n'importe quel contrôle afin de répondre à une tâche spécifique. Remplaçant un périphérique physique bardé de boutons et de capteurs, les systèmes tactiles dématérialisent ces actionneurs physiques en les

¹⁵³ NAKATANI Lloyd H. et ROHRlich John A. *Soft machines: A philosophy of user-computer interface design*. In *Proceedings of the SIGCHI conference on Human Factors in Computing System* (CHI '83). ACM, New York, USA, 1983, p. 19-23

¹⁵⁴ « *Soft-controls* ». In NAKATANI Lloyd H. et ROHRlich John A. *Soft machines: A philosophy of user-computer interface design*. In *Proceedings of the SIGCHI conference on Human Factors in Computing System* (CHI '83). ACM, New York, USA, 1983, p. 19-23. Traduction Libre : COQUELIN Marine et MORIN Rudy.

¹⁵⁵ NAKATANI Lloyd H. et ROHRlich John A. *Soft machines: A philosophy of user-computer interface design*. In *Proceedings of the SIGCHI conference on Human Factors in Computing System* (CHI '83). ACM, New York, USA, 1983, p. 19-23

¹⁵⁶ « *Less is more* » BUXTON Bill. *A touching story : A personal perspective on the history of touch interfaces past and future*. In *Society for Information Display (SID) Symposium Digest of Technical Papers* vol. 41, 2010, pp. 444-448. Traduction Libre : COQUELIN Marine et MORIN Rudy.

représentant à l'écran sous la forme la plus appropriée ou parfois de manière transparente. Cela confère à l'interface davantage de flexibilité qu'un périphérique d'entrée physique.

L'apparition des interfaces tactiles multi-points a cependant ouvert la voie à des opérations de contrôle plus complexes. Là où les interfaces tactiles à contact unique n'autorisent que des actions indépendantes dans le temps, les interfaces tactiles multi-points autorisent des interactions simultanées dans l'espace, de manière indépendante ou en participant à la même opération de contrôle. Là où les unes permettent de manipuler l'interface par une gestuelle limitée, les autres permettent l'utilisation d'une gestuelle plus complexe, répondant aux fonctions ergotiques, symboliques, pictographiques, etc. D'un côté, la manipulation se fait avec un doigt ou un stylet, de l'autre, la manipulation peut se faire avec un doigt ou plusieurs, avec la main ou les deux. Les interfaces tactiles multi-points font donc appel à des modalités d'interaction nouvelles, étroitement liées à la technologie tactile utilisée.

Ce que nous proposons ici est de faire l'inventaire de principes théoriques, de règles et de propriétés qui vont constituer un modèle d'interaction propre à ce type d'interface. Il s'agit principalement de décrire les modalités d'interaction en faisant émerger leurs spécificités et en caractérisant les relations entre l'utilisateur et l'interface. Selon Michel Beaudouin-Lafon¹⁵⁷, un modèle d'interaction permet de définir le « *look and feel* » des interactions, c'est à dire comment l'utilisateur va percevoir et ressentir ces interactions. Cela consiste à caractériser l'interface de manière à permettre au designer d'interaction de créer ou de comprendre les aspects de son design, des éléments qui la composent (*look*) et de bénéficier d'un espace de compréhension pour appliquer ou saisir les comportements interactifs (*feel*) des objets graphiques qui la composent.

Michel Beaudouin-Lafon¹⁵⁸ fait la description d'un modèle d'interaction en distinguant les propriétés suivantes :

- Un pouvoir descriptif (de classification). C'est à dire la capacité de décrire et de classer une interface existante.

¹⁵⁷ BEAUDOUIN-LAFON Michel. *Instrumental interaction : an interaction model for designing post-WIMP user interfaces*. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (CHI '00). ACM, New York, USA, pp. 446-453

¹⁵⁸ *ibid.*

- Un pouvoir génératif. Il s'agit de la capacité d'aider un designer d'interface à créer de nouveaux designs d'interaction.
- Un pouvoir comparatif. C'est la capacité du modèle d'interaction à fournir les éléments d'analyse afin d'évaluer un design d'interface.

Dans ce chapitre, nous déclinons notre modèle selon trois aspects déterminants. D'une part, nous nous attarderons sur les modalités du geste dans sa relation à l'image, son exécution, sa dynamique et ses caractéristiques. Nous décrirons ensuite les modalités des interactions en insistant sur les spécificités et les contraintes de leur exécution. Enfin, nous porterons notre attention sur le rôle des attributs physiques et technologiques des dispositifs dans la conception d'interface utilisateur.

1. Interface utilisateur naturelle

1.1. Terminologie

Les recherches menées sur l'interaction gestuelle dans le domaine des IHM et plus particulièrement les travaux sur la réalité augmentée et la virtualité augmentée considèrent le geste comme une modalité d'interaction avec le système permettant de produire des formes d'interactions dites plus naturelles. Cette propriété concerne également les interfaces multi-tactiles qui s'inscrivent dans une heuristique d'interaction couramment appelée interface utilisateur naturelle (*Natural User Interface* ou NUI). Cette appellation, introduite par Christian Moore en 2006, s'est progressivement déplacée de la sphère de la recherche au milieu industriel ou celui des journaux grand public. De nombreuses sociétés se sont tournées vers le développement et la mise au point de produits ou surcouches logicielles qu'ils mettent à la vente en basant leur communication sur l'innovation et les promesses contenues dans la définition de ces termes. C'est le cas par exemple des démonstrations vidéo¹⁵⁹ de la *Microsoft Surface* ou d'entreprises comme *Nuiteq*¹⁶⁰ qui profitent de la dénomination NUI. Au cours des chapitres précédents, nous avons mis en

¹⁵⁹ Disponible sur :

<http://www.microsoft.com/showcase/en/us/details/0fad0557-4629-4622-a16b-b3443eedf42a>

¹⁶⁰ Disponible sur : <http://www.nuiteq.com>

lumière l'histoire de ces technologies et l'envergure des travaux de recherche scientifiques qui ont accompagnés leur évolution. La plupart des développeurs et des designers n'ont pas toujours idée de la complexité et des enjeux que peut soulever la création d'interface gestuelles et cela conduit au développement de logiciels ou de matériels dont les caractéristiques sont complètement en décalage par rapport aux concepts et aux idées amenées par cinquante ans d'histoire de la recherche en IHM. C'est en partie ce constat qui m'a amené, dans le cadre du contrat CIFRE, à conduire mon projet de recherche de doctorat sur la conception et le développement d'un intergiciel multi-tactile (cf. Chapitre VI) qui répondrait davantage à nos besoins et qui s'inscrirait dans une continuité de la recherche.

La confusion qui entoure la définition de la dénomination « interface utilisateur naturelle » provient essentiellement de l'implication du terme « naturel », bien trop vague pour saisir toute la portée des concepts mis en jeu dans les interactions tactiles et gestuelles. La définition du terme naturel donnée par le *Petit Robert* comme ce « *qui appartient à la nature d'un être, d'une chose* » ne peut que nous suggérer de creuser davantage les caractères du geste en interaction homme-machine. Il existe en anglais le terme *naturalness* qui ne possède pas réellement d'équivalent en français puisque naturalité renvoie davantage à la notion de territoire ou de nationalité. Le terme naturel est souvent associé à celui d'intuitif qui n'éclaircit pas forcément davantage les concepts sous-jacents aux NUI. Il est possible d'aborder la définition du terme naturel en considérant qu'il décrit, dans le cas des IHM, les gestes de la main correspondant à un large spectre d'actions issues d'aptitudes manuelles existantes ou impliquant une faible charge cognitive. De manière générale, nous pourrions rapprocher cette définition de celle avancée par Bill Buxton¹⁶¹ qui attribue le caractère naturel à une interface si elle « *exploite les compétences que nous avons acquises durant notre vie en vivant en ce monde.*¹⁶² »

Dans un article récent¹⁶³, Ron George et Joshua Blake proposent de nouvelles métaphores pour décrire la nature des interfaces utilisateur naturelles : les objets, les containers, les gestes et les manipulations (*Objects, Containers, Gesture, Manipulation* ou OCGM). De même que le paradigme WIMP utilise la métaphore du bureau ou des dossiers, il est possible

¹⁶¹ BUXTON Bill. *CES 2010 : NUI with Bill Buxton*. Larry Larsen et Bill Buxton. Channel9. Disponible sur : <http://channel9.msdn.com/Blogs/LarryLarsen/CES-2010-NUI-with-Bill-Buxton>

¹⁶² « [...] exploits skills that we have acquired through a lifetime of living in this world ». In *ibid*.

¹⁶³ GEORGE Ron et BLAKE Joshua. *Objects, Containers, Gestures, and Manipulations: universal foundational metaphors of natural user interfaces*. In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '10)*, Atlanta, Georgia, USA, 2010

de décrire les interactions tactiles gestuelles en rapprochant leurs modalités à des formes du réel. D'une part, les *objets* correspondent à des unités de données ou une forme de contenu. Les *containers* représentent les relations spatiales ou fonctionnelles entre les différents contenus. La gestuelle concerne les interactions discrètes ou nécessitant un apprentissage tandis que les *manipulations* correspondent aux gestes continus d'interaction avec l'environnement : les gestes ergotiques. Ron George distingue les gestes de manipulation des autres car ils sont « *simples, intuitifs, et d'une certaine manière enveloppés d'une métaphore* ». Les autres seraient « *complexes, appris, non physiques et super-naturels.* »

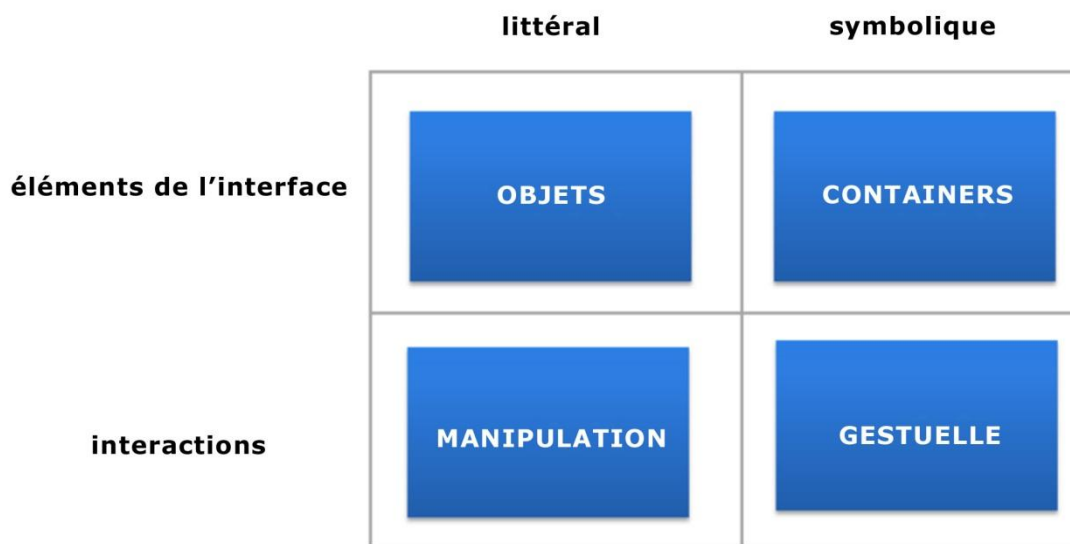


Figure 51 - Classification proposée par Ron George et Joshua Blake. In [GEORGE et BLAKE, 2010].

George et Blake avancent que la limitation des interfaces utilisant la métaphore WIMP ne permet pas de réduire la charge cognitive des interactions en dessous d'un certain seuil. Afin d'y parvenir, les auteurs suggèrent d' « *identifier les différentes habiletés innées du cerveau humain et de concevoir des interfaces qui prennent avantage de ces compétences naturelles dans un contexte approprié.*¹⁶⁴ » En mettant en perspective les travaux de Jean Piaget sur le développement de l'intelligence sensori-motrice dans le contexte des IHM¹⁶⁵, ils appuient leur hypothèse en comparant la charge cognitive nécessaire pour comprendre et utiliser des interfaces WIMP à celle nécessaire pour utiliser les interfaces utilisateur naturelles de type

¹⁶⁴ « [...] identify the various innate abilities of the human brain and design interfaces which take advantage of those natural skills in the appropriate contexts ». In *ibid.* Traduction libre : COQUELIN Marine et MORIN Rudy.

¹⁶⁵ **PIAGET Jean.** *L'épistémologie génétique.* PUF, 2005, 126 p. (Coll. Que sais-je ?)

OCGM. Il apparaît que les compétences cognitives nécessaires aux NUI apparaissent plus tôt dans le développement de l'enfant et sont considérées comme « *innées et naturelles* » par les auteurs contrairement aux interfaces WIMP qui nécessitent des compétences davantage spécialisées et qui n'apparaissent que plus tardivement.

Les interfaces tactiles sont corpo-centrées et tendent à définir un espace d'action dont les coordonnées du système correspondent aux coordonnées de l'utilisateur. Nous pouvons alors ajouter à l'acception du terme naturel, comme élément de compréhension, la capacité d'une interface à s'inscrire dans son espace d'action physique. Les interfaces utilisateur naturelles sont davantage un concept qu'un paradigme d'interaction, un ensemble de principes tels ceux décrits par le concept de manipulation directe par Ivan Sutherland qui donnent une direction théorique à la recherche en IHM. Au contraire, les notions d'OCGM se prêtent particulièrement bien pour décrire les interactions utilisateur tactiles et gestuelles et c'est le paradigme que nous retiendrons pour la définition de notre modèle d'interaction et le développement de l'intergiciel. L'acronyme se prononce ocam et fait référence également au rasoir d'Occam, principe de raisonnement selon lequel « *les multiples ne doivent pas être utilisés sans nécessité*¹⁶⁶ » ou dans une interprétation plus moderne : « *les hypothèses les plus simples sont les plus vraisemblables* » Il est alors possible dans un premier temps de rapprocher le caractère naturel d'une interface à la notion de simplicité.

1.2. Métaphore du réel

L'homme interagit avec l'environnement grâce à la richesse de ses moyens de perception, de communication et d'action dans le monde physique. Il bénéficie de la vue, de l'ouïe, du goût, de l'odorat et des facultés de perception tactilo-kinesthésiques et proprioceptives qui lui permettent d'appréhender l'espace physique des objets. Nous utilisons les gestes et la richesse de nos prédispositions anatomiques pour communiquer et manipuler les objets. En comparaison, les interactions homme-machine traditionnelles nous semblent bien limitées. Le paradigme des interfaces utilisateur naturelles propose de dépasser cette limitation en explorant de nouvelles modalités d'interaction en exploitant les canaux de communication et de manipulation que nous utilisons naturellement et quotidiennement. C'est le cas des interfaces utilisateur tangibles qui créent un espace de communication et d'action mixte et

¹⁶⁶ « *pluralitas non est ponenda sine necessitate* »

celui des interfaces haptiques ou multimodales qui donnent au corps la possibilité de s'exprimer dans l'espace réel. Les surfaces interactives multi-tactiles s'inscrivent également dans cette heuristique d'interactions en accordant au geste un espace d'action (manipulation) et d'expression (gestuelle) plus étendu.

En 2008, Jacob *et al.*¹⁶⁷ introduisent le concept d'« *interaction basée sur le réel* » (*Reality-based interaction* ou RBI) qui, selon les auteurs, permettrait d'analyser et d'avoir une meilleure compréhension du rapprochement des interactions homme-machines à celles du réel. L'emploi du terme « réel » peut poser problème car il renvoie à différentes interprétations selon le contexte social et culturel, c'est pourquoi nous précisons qu'il fait référence au monde physique des objets en opposition avec le monde numérique des machines, des processus et des périphériques. Afin de proposer un cadre conceptuel au concept des RBI, Jacob *et al.* distinguent quatre thèmes déterminants dans l'émergence de ces nouvelles formes d'interaction :

- La physique naïve (*naïve physics*). La physique naïve correspond à la faculté informelle humaine de percevoir les principes physiques élémentaires, c'est-à-dire la connaissance du comportement des objets dans le monde physique. Nous avons tous, d'une certaine manière, plus ou moins consciente, des notions de gravité, de vitesse, de forces de frottement, d'inertie et d'échelle. Dans le domaine de l'intelligence artificielle, la physique naïve fait référence aux tentatives pour décrire le monde physique, par des algorithmes, davantage du point de vue de l'homme que de celui des sciences physiques. Dans le cas des interfaces tactiles, si la manipulation d'objets graphiques répond à une certaine logique physique, l'utilisateur aura instinctivement une meilleure compréhension des modalités d'interaction. Le déplacement d'une image peut être soumis par exemple à l'illusion d'une gravité et de forces de frottements qui vont lui conférer son accélération, sa vitesse et son inertie. Les objets numériques réagissent à la manipulation comme le feraient des objets physiques, ce qui renforce l'impression de contrôle chez l'utilisateur.

¹⁶⁷ JACOB Robert J.K., GIROUARD Audrey, HIRSHFIELD Leanne M., HORN Michael S., SHAER Orit, SOLOVEY Erin Treacy, et ZIGELBAUM Jamie. *Reality-based interaction : a framework for post-WIMP interfaces*. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems* (CHI '08). ACM, New York, USA, 2008, pp. 201-210

- La conscience du corps (*Body Awareness*). Il s'agit principalement des facultés naturelles de l'homme à connaître son corps, indépendamment de son environnement. C'est la conscience que nous avons de notre motricité, des positions relatives de nos membres (proprioception) et des degrés de liberté des articulations qui nous donnent la connaissance de notre zone d'actions et de mouvements. C'est également notre capacité à coordonner plusieurs mouvements afin de réaliser des tâches d'apparences triviales comme marcher ou conduire une voiture et qui pourtant font appel à la coordination de plusieurs centaines de muscles et articulations. Dans le cadre des interactions multi-tactiles, cette faculté peut être mise à profit, notamment lorsque l'interface propose des interactions bi-manuelles. Les deux mains peuvent agir simultanément à l'accomplissement d'une même tâche (un zoom à deux mains) ou au contraire s'acquitter de tâches individuelles (déplacer deux images indépendamment).

- La conscience de l'environnement (*Environment Awareness*). Nous avons tous la capacité de nous situer dans notre environnement spatial, composé d'objets arrangés dans un espace en trois dimensions. Cette faculté nous donne les dispositions nécessaires à notre orientation et aux notions d'échelle et de distance dont nous nous servons par exemple pour nous diriger. Cette faculté peut s'avérer utile dans le cas d'univers virtuels en trois dimensions où l'utilisateur doit intégrer la disposition spatiale des objets graphiques et les manipuler. Naviguer dans un univers virtuel en trois dimensions est également simplifié par notre prédisposition à évoluer dans l'espace réel. Nous avons développé la faculté de manipuler les objets dans leur environnement, en les saisissant (préhension) ou en appliquant une force afin de modifier leurs propriétés (forme, position, etc.). C'est cette prédisposition qui est exploitée lorsque l'utilisateur fait défiler des images en les déplaçant d'un point à un autre, comme s'il s'agissait de photos posées sur une table.

- La conscience sociale (*Social Awareness*). L'homme a conscience de la présence des autres individus et il a développé des aptitudes pour communiquer et interagir dans un contexte social. Dans le cadre des interfaces tactiles, deux possibilités sont offertes au designer d'interaction. D'une part, il peut profiter des gestes sémiotiques existant comme le geste déictique qui indique une direction et peut représenter la

cible d'un déplacement d'objet graphique ou créer de nouvelles unités sémantiques, ou lemmes, dans un champ lexical propre à son application comme poser dix doigts sur la surface pour donner l'indication au système de faire apparaître un clavier virtuel. Dans le contexte social, notre rapport à l'autre est aussi défini par les interactions physiques qui interviennent au travail ou de manière générale comme donner et recevoir un objet ou s'entraider pour la réalisation d'une même tâche. Analyser les modalités de ces relations peut aider à concevoir des systèmes multiutilisateurs ou des plateformes collaboratives de travail où la présence de l'autre est soit physique (autour d'une table interactive par exemple) ou représentée par un avatar (interactions en réseau).

1.3. Limitations de la métaphore du réel

Au cours du chapitre précédent, nous avons décrit le cadre conceptuel de la métaphore du réel pour l'émergence de nouvelles modalités d'interaction. Jacob *et al.* font le constat que de nombreux systèmes émergents comme les interfaces utilisateur tangibles, tactiles ou haptiques, peuvent être analysés à travers le prisme de la classification qu'ils proposent. Les auteurs partent du postulat qu'en créant des interactions basées sur les connaissances et les facultés de l'homme dans son environnement naturel, l'effort mental nécessaire pour interagir avec le système, et les temps d'apprentissage, sont considérablement réduits car les utilisateurs possèdent déjà les compétences nécessaires. Les correspondances entre le monde physique et numérique augmenteraient l'efficacité de tels systèmes et « [...] encouragerait l'improvisation et l'exploration car les utilisateurs n'ont pas à apprendre des compétences spécifiques à l'interface.¹⁶⁸ »

Toutefois, nous serions en droit de nous demander si cette direction n'implique pas quelques limites techniques et conceptuelles. D'une part, nous pouvons douter de l'efficacité des systèmes actuels à analyser et interpréter toute l'étendue de la signification du geste, toute sa granularité. D'autre part, est-il réellement opportun de créer des

¹⁶⁸ « [...] encourage improvisation and exploration because users do not need to learn interface-specific skills. » JACOB Robert J.K., GIROUARD Audrey, HIRSHFIELD Leanne M., HORN Michael S., SHAER Orit, SOLOVEY Erin Treacy, et ZIGELBAUM Jamie. *Reality-based interaction : a framework for post-WIMP interfaces*. In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems (CHI '08)*. ACM, New York, USA, 2008, pp. 201-210. Traduction libre : COQUELIN Marine et MORIN Rudy.

interactions naturelles pour l'ensemble des commandes et des manipulations du système sachant que le numérique implique justement la création d'artefacts ou de commandes abstraites qui n'existent pas dans le réel. Il est alors nécessaire d'évaluer les coûts les avantages de cette modalité d'interaction afin de trouver un juste équilibre entre réalité et virtualité. Pour répondre à cette problématique, Jacob *et al.* proposent de mettre en balance le réalisme des interaction avec l'usabilité de l'interface :

- *Réalisme et fonctionnalité.* Certains systèmes complexes peuvent entraîner la multiplicité de commandes et de fonctions. Cependant, la complexité d'un système peut considérablement diminuer l'efficacité de l'utilisateur si l'interface ne dispose pas de macro-commandes ou de méta-outils pour simplifier les interactions. A vouloir systématiquement imiter les interactions du réel, les interactions peuvent perdre en efficacité. Par exemple, la manipulation d'images graphiques représentées sur une surface interactive comme elles le seraient sur une table physique permet de les manipuler avec une certaine simplicité mais devoir toutes les organiser hiérarchiquement si elles sont en grand nombre deviendrait vite plus compliqué.
- *Réalisme et efficacité.* Certaines tâches sont réalisées plus rapidement si les moyens de les réaliser sont plus efficaces. L'édition vidéo est un domaine où les raccourcis clavier sont largement utilisés. Revenir à la métaphore du réel, en coupant la pellicule et en la déplaçant avec le geste prendra certainement plus de temps que si l'opération était effectuée à la souris par exemple, l'amplitude des mouvements étant largement supérieure sur la surface interactive.
- *Réalisme et versatilité.* En spécialisant l'interface graphique, elle gagne en réalisme mais réduit la variété de tâches qu'il est possible de réaliser.
- *Réalisme et ergonomie.* Certains dispositifs comme les tables interactives ou les murs d'images tactiles imposent des mouvements amples et répétés qui peuvent nuire à l'ergonomie du système en causant fatigue et stress musculaire. C'est le cas lorsque l'on garde le bras trop longtemps tendu vers l'avant, en contact avec un écran vertical.

- *Réalisme et accessibilité.* Dans certains cas, les interactions trop réalistes peuvent causer un problème d'accessibilité aux personnes âgées ou physiquement diminuées. Il faut considérer à l'avance le contexte d'utilisation de l'interface et préférer l'utilisation d'outils plus adaptés.
- *Réalisme et praticité.* Il est nécessaire d'évaluer des aspects plus pratiques comme le coût, les contraintes techniques, d'espace, de durabilité, de consommation et d'impact sur l'environnement que supposent ces systèmes interactifs.

2. Modélisations temporelles du geste

2.1. Forme canonique du geste

En considérant le geste comme un processus dynamique, qui s'inscrit dans le temps, il semble nécessaire de définir les caractéristiques temporelles de sa production et de déterminer son couplage avec l'interface. Les contributions scientifiques dans le domaine de la psychologie et de la linguistique nous donnent un cadre théorique général pour la compréhension du procès de la réalisation du geste. Il s'agit d'analyser la forme canonique du geste et le cours kinésique du mouvement en segments temporels afin de déterminer le moment de la « *phrase gestuelle*¹⁶⁹ » qui intéresse notre modèle d'interaction. Kendon distingue trois phases principales : la préparation, la phase nucléaire (apogée ou trait), et la rétractation¹⁷⁰.

- La phase de préparation correspond à l'impulsion préparatoire qui met la main en mouvement après une phase de repos.

¹⁶⁹ KENDON Adam. *Current issues in the study of gesture*. In *Biological Foundations of Gesture Hillsdale*. New Jersey, Lawrence Erlbaum Associates, 1986, pp. 23-47

¹⁷⁰ « *Preparation, nucleus (peak or stroke) and retraction* » In KENDON Adam. *Current issues in the study of gesture*. In *Biological Foundations of Gesture Hillsdale*. New Jersey, Lawrence Erlbaum Associates, 1986, pp. 23-47. Traduction libre : COQUELIN Marine et MORIN Rudy.

- La phase « nucléaire » (*nucleus*) correspond au geste porteur de sens car il possède « une forme définie et des propriétés dynamiques avancées¹⁷¹ ». Il s'agit du moment où le geste prend sa forme la plus substantielle, où le geste exprime toute son intentionnalité. Au cours de cette phase, la gesticulation peut être simple ou complexe selon le nombre de mouvements (*stokes*) décrit par la phrase gestuelle. De plus, le geste nucléaire peut être statique (les mains ne décrivent aucun mouvement) ou dynamique (mouvements internes de la main et/ou mouvements du bras).
- La phase de rétraction correspond au retour de la main en position de repos ou à son repositionnement pour la réalisation d'un nouveau geste et dans ce cas, la rétractation n'est que partielle.

Les phases de préparation et de rétractation sont généralement relativement rapides alors que la phase nucléaire est caractérisée par un mouvement plus complexe et plus long. Il convient également de caractériser la « manualité » (*handedness*) du geste, à savoir s'il est réalisé à une main (dominante/non dominante) ou à deux mains.

2.2. Etats du système

Aux mouvements de la main correspond différents états de la main qu'il est possible de définir grâce aux étapes définies par Dean Rubine¹⁷² qui distingue quatre étapes principales de l'interaction avec le système :

- L'attente (*WAIT state*) correspond à l'état d'un système qui attend le début du geste de l'utilisateur.
- L'acquisition (*COLLECT state*) des données qui commence lorsque le geste débute. Durant cette phase, le système capture l'ensemble des données recueillies par le périphérique d'entrée avant de l'analyser et de l'interpréter. Dans le cas des interfaces tactiles, nous pouvons associer le début de cette phase lorsque l'utilisateur touche l'écran alors qu'aucun autre contact n'est présent.

¹⁷¹ KENDON Adam. *Current issues in the study of gesture*. In *Biological Foundations of Gesture Hillsdale*. New Jersey, Lawrence Erlbaum Associates, 1986, pp. 23-47

¹⁷² RUBINE Dean Harris. *The Automatic Recognition of Gestures*. Thèse. Carnegie Mellon University, Pittsburgh, USA 1991

- La manipulation (*Manipulate state*) intervient lorsque le geste a été analysé et classifié. Ce moment est soit explicite s'il correspond à un événement généré par le périphérique ou implicite si la classification est établie selon des paramètres d'interaction (la pose prolongé d'un geste dont la classification est réalisée au bout d'un temps donné). Enfin, la classification du geste peut également se faire alors que le geste lui-même n'est pas terminé et lorsque le système a collecté suffisamment d'informations pour l'identifier sans ambiguïté. Lorsque cette phase commence, le système est en mesure de fournir le feedback correspondant au geste identifié.
- L'exécution (*EXECUTE state*) intervient lorsque l'utilisateur a terminé l'interaction. Dans le cas des interfaces tactiles, c'est le moment où l'utilisateur retire ses doigts de l'écran ou de l'objet de la manipulation. Le système revient alors à son état d'attente prêt pour analyser un nouveau geste.

Le schéma ci-dessous permet de mettre en relation les étapes de la phrase gestuelle et les états d'un système tactile. Dans ce modèle, j'écarte donc de l'analyse système les mouvements effectués lors de la préparation du geste. Nous verrons que le choix, a priori, de ces modalités de couplage entre le geste et le système sera revu et précisé lors des phases de tests de l'environnement de développement.

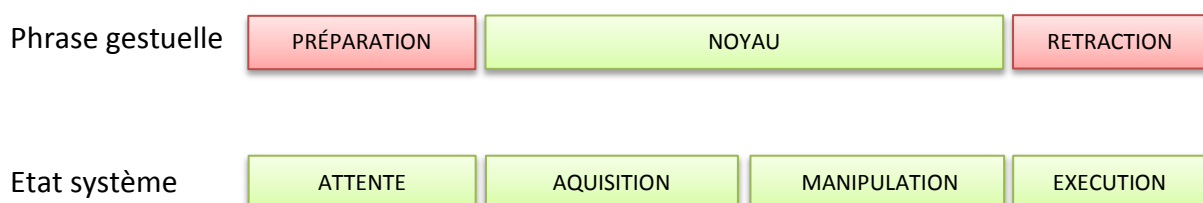


Figure 52 - Représentation du couplage de la phrase gestuelle avec les états du système.

2.3. Discret / continu

Il s'agit de définir le flux du geste selon son couplage temporel avec le système. Il peut être discret si la commande attendue par le geste intervient lorsque celui-ci est terminé. C'est le cas lorsque l'utilisateur trace à l'écran le signe de point d'interrogation afin d'ouvrir une fenêtre d'aide. Le système attend la fin du geste avant d'appliquer un algorithme d'interprétation de mouvement. Le flux peut être continu si le système analyse le geste dès le moment où la main touche l'écran et associe le geste à la modification progressive d'éléments graphiques. C'est le cas d'un simple glisser / déposer où le système déplace l'objet en temps réel en fonction de la position du ou des points de contact.

3. Modélisation spatiale du geste

3.1. Interface multiplexée dans l'espace

Selon George Fitzmaurice¹⁷³, les périphériques d'entrée utilisateur peuvent appartenir soit aux dispositifs à multiplexage temporel, soit aux dispositifs à multiplexage spatial. Dans le cas des interfaces multiplexées dans l'espace, chaque fonction est contrôlée par un transducteur unique et indépendant dans l'espace. A l'inverse, les interfaces multiplexées dans le temps ne requièrent qu'un seul périphérique pour contrôler différentes fonctions en des temps successifs. Fitzmaurice décrit les interfaces WIMP traditionnelles comme multiplexées dans le temps car la souris contrôle diverses fonctions comme la sélection, la navigation, l'activation de boutons, de manière consécutive. Bien que le clavier soit utilisable simultanément avec la souris, les touches de raccourcis sont principalement activées afin de soutenir l'action réalisée par la souris et se partagent ainsi la réalisation d'une même tâche. Il existe donc une dichotomie entre l'espace d'action de l'interface qui est généralement multiplexé dans l'espace (plusieurs éléments de contrôle graphiques sont présents simultanément) et les périphériques d'entrée qui sont multiplexés dans le temps. Ce type d'interface a pour conséquence des interactions séquentielles et principalement exclusives.

¹⁷³ FITZMAURICE George W. *Graspable User Interfaces*. Thèse, University of Toronto, Toronto, Canada, Canada. Sous la direction de William Buxton, 1996

Dans le cas des interfaces tactiles multi-points, le périphérique d'entrée supporte plusieurs points de contacts simultanément et chacun d'eux peut intervenir sur un contrôle indépendant. Plusieurs cas de figure sont possibles :

- Une main contrôle une fonction unique sur un objet unique.
- Une main contrôle plusieurs fonctions sur un objet unique (exemple d'un zoom et d'une rotation simultanée sur un objet).
- Une main contrôle plusieurs fonctions sur des objets différents (zone limitée car chaque doigt est rattaché à la même main et dépend de sa zone d'action).
- Deux mains participent à la même tâche (zoom sur une image réalisé à deux mains par exemple).
- Deux mains participent à des tâches différentes (déplacer ou zoomer sur des images indépendantes dans l'espace, la zone d'action étant limitée à l'envergure des bras).
- Deux mains contrôlent des fonctions différentes mais partagent la même tâche (une main active un contrôle pendant que l'autre ajuste un paramètre).

Enfin, il nous est possible de classifier le geste selon sa « dynamicité » spatiale. Au regard de la classification proposée par Harling et Edwards¹⁷⁴ issue de leur contribution sur la langue des signes, nous pouvons également décliner la spatialité du geste selon sa dynamique externe ou interne. Quatre configurations sont alors possible :

- *Pose statique, position spatiale statique.* Ex : l'utilisateur pose ses dix doigts sur l'écran sans les bouger, afin de faire apparaître un clavier virtuel.
- *Pose dynamique, position spatiale statique.* Ex : l'utilisateur pose le pouce et l'index sur un objet graphique et l'agrandit en écartant les doigts (pose dynamique) sans modifier la position de sa main (position spatiale statique).
- *Pose statique, position spatiale dynamique.* Ex : l'utilisateur pose un ou plusieurs doigts sur un objet graphique et le déplace en bougeant la main (position spatiale dynamique) sans changer la configuration interne de la position des doigts (pose statique).

¹⁷⁴ HARLING Philip A. et EDWARDS Alistair D. N. *Hand Tension as a Gesture Segmentation Cue. In Proceedings of gesture workshop on progress in gestural interaction*, Springer-Verlag, London, UK, 1996, pp. 75-88

- *Pose dynamique, position spatiale dynamique.* Ex : l'utilisateur pose le pouce et l'index sur un objet graphique et modifie de manière simultanée sa taille et sa position, d'une part en écartant les doigts (pose dynamique) et d'autre part en déplaçant la main sur la surface.

3.2. Degrés de liberté

La richesse des interactions est largement conditionnée par la richesse et/ou le nombre de degrés de liberté du périphérique d'entrée utilisateur. Dans le cas des interfaces tactiles, l'instrument d'interaction physique n'est autre que la main de l'utilisateur qui a profité de toute l'histoire de l'humanité pour développer son habileté, sa dextérité et son adaptabilité. Je propose ici de nous attarder sur les spécificités et les caractéristiques de la main comme instrument d'interaction.

J.R. Napier¹⁷⁵ distingue deux types de mouvements. D'une part, les mouvements préhensifs sont ceux où l'objet est saisi ou tenu intégralement dans la main. D'autre part, les mouvements non préhensifs où l'objet n'est pas saisi par la main mais peuvent être manipulés en les poussant ou en appliquant une force des doigts de la main. Les mouvements de la main sur une surface interactive correspondent à cette deuxième catégorie en ce sens où la manipulation des objets graphiques se fait en le touchant du doigt ou de la main sans véritablement saisir l'objet puisque l'interaction se fait sur une surface en deux dimensions.

Toutefois, les schémas réalisés par J.R. Napier révèlent toute la complexité et la richesse anatomique des quelques 23 degrés de liberté de la main (en comptant uniquement les degrés de liberté liés aux articulations) qui lui confère sa flexibilité, ses prédispositions et ses contraintes comme instrument d'interaction tactile. Les os du squelette de la main sont reliés entre eux par des articulations dont les degrés de liberté varient de un à trois. L'index, le majeur, l'annulaire et l'auriculaire, appelé également doigts longs, possèdent trois articulations. Le pouce ne possède que deux articulations mais dispose d'une mobilité plus importante.

¹⁷⁵ NAPIER John R. *The prehensile movements of the human hand*. In *Surger* 38-B(4), 902-913. 1956. Disponible ici : <http://www.ncbi.nlm.nih.gov/pubmed/13376678>

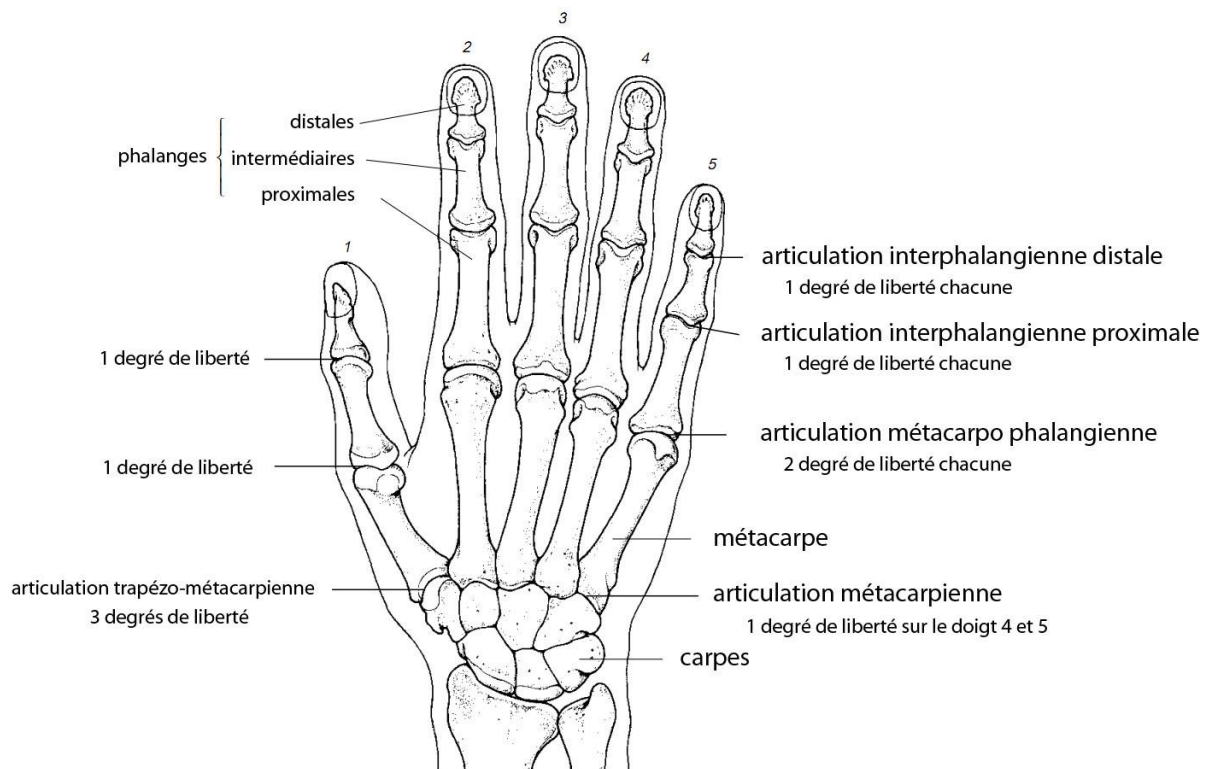


Figure 53 - Articulations de la main. Ce schéma présente les 17 articulations de la main fournissant 23 degrés de libertés. Reproduction. In [Napier, 1980].

De plus, la main bénéficie des mouvements de flexion et d'abduction du poignet ainsi que des mouvements complexes de pronation et de supination qui permettent d'effectuer une rotation de la main. Sa complexité articulaire permet de multiples configurations et de multiples mouvements internes. Le schéma suivant nous donne une représentation générale des différents mouvements d'abduction, d'extension et de flexion des doigts de la main.

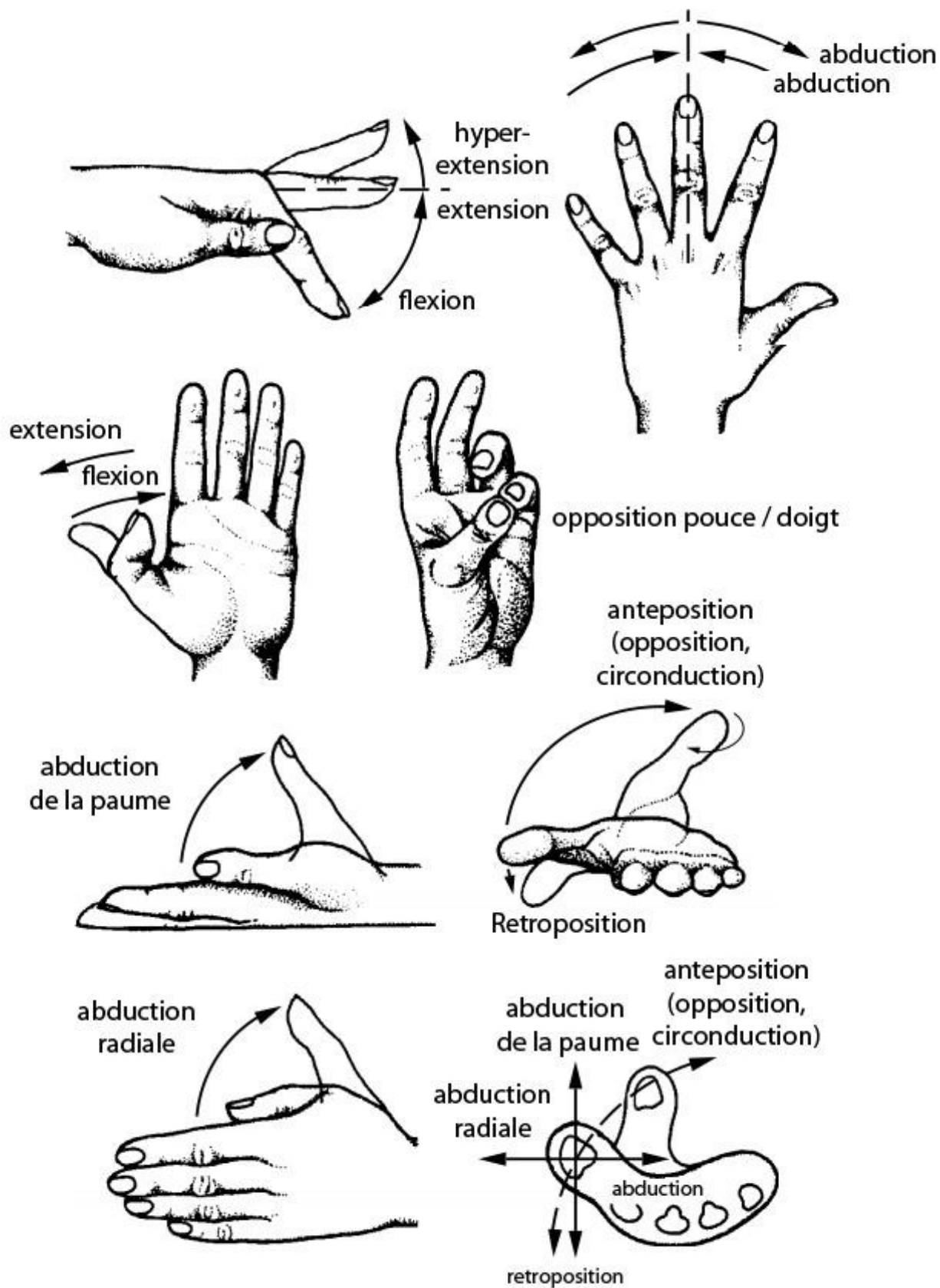


Figure 54 - Mouvements articulaires de la main. Schéma de l'American Society for Surgery of Hand, 1978.
In [Sturman, 1992].

3.3. Le degré d'indirection

Le degré d'indirection¹⁷⁶ est la mesure en deux dimensions de la compensation spatiale et temporelle générée par un instrument d'interaction. Le degré spatial est la distance entre la partie logique de l'instrument et l'objet d'intérêt à l'écran. Dans le cas des interfaces WIMP traditionnelles, le déplacement d'une fenêtre implique une compensation spatiale minimale car l'objet d'intérêt est situé sous l'instrument d'interaction (curseur de la souris). Au contraire, le défilement d'une fenêtre en déplaçant la barre de défilement implique une compensation spatiale plus importante étant donné que l'instrument se trouve à une certaine distance de l'objet d'intérêt. Dans le cas des interfaces tactiles, une faible compensation spatiale est préférable car les deux parties, logique et physique de l'instrument (le doigt ou la main) sont spatialement situées au même endroit que l'objet d'intérêt. Cette proximité spatiale crée les conditions d'une logique de cause à effet qui unit les actions de l'utilisateur aux réponses du système de façon cohérente.

La compensation temporelle correspond au temps de réaction de la transformation de l'objet par rapport à l'action physique sur l'instrument d'interaction. Dans le cas des interfaces traditionnelles, la modification d'une propriété de l'objet d'intérêt requiert la validation de cette modification pour que son effet soit appliqué sur l'objet. Une faible compensation est préférable dans le sens où l'utilisateur a un retour immédiat de l'action qu'il est en train de réaliser. Dans le cas des interfaces utilisateur multi-tactiles, la compensation temporelle est souvent étroitement liée à la latence générale du système, à laquelle s'ajoutent les temps d'acquisition et de traitement des données tactiles.

3.4. Le degré d'intégration

Le degré d'intégration¹⁷⁷ mesure le rapport entre le nombre de degrés de liberté proposé par la partie logique de l'instrument et le nombre de degrés de liberté mesuré par le périphérique d'entrée utilisateur. Dans les interfaces traditionnelles, le nombre de degrés de liberté proposé par la partie logique de l'instrument est généralement situé entre 1 et 3

¹⁷⁶ BEAUDOUIN-LAFON Michel. *Interaction Instrumentale : de la manipulation directe à la réalité augmentée*. IHM'97, 9èmes Journées sur l'Ingénierie de l'Interaction Homme-Machine. Cépadués Editions, 1997, pp. 97-104

¹⁷⁷ *ibid.*

puisque les propriétés modifiables de l'instrument ou de l'objet sont généralement limitées à ce nombre.

Déplacer une vue utilisateur par l'intermédiaire d'une barre de défilement verticale implique un seul degré de liberté pour l'instrument (axe des y) d'interaction alors que la souris dispose de deux degrés de liberté (x et y). Le degré d'intégration de cet instrument est donc de $1/2$. Afin d'ajuster la vue sur l'axe des x, il faudra refaire la manipulation sur la barre de défilement horizontale rendant l'opération plus longue moins efficace. Les modifications des propriétés spatiales d'un objet graphique par glisser / déposer, réalisées grâce à l'instrument de déplacement représenté par une main, possède deux degrés de liberté, x et y, son degré d'intégration et alors de $2/2$. Le contrôle d'objets en trois dimensions dans un logiciel de modélisation 3D implique généralement des instruments qui ont un degré d'intégration de $3/2$ puisque l'instrument actif permet par exemple de modifier les trois positions de l'objet dans l'espace. La représentation graphique de l'instrument joue alors un rôle important pour rendre plus explicite les moyens d'action sur l'objet.

L'efficacité du système, le coût de la tâche effectuée et son niveau de complexité sont étroitement liés au degré d'intégration de l'interface. Toutefois, nous serions en droit de nous demander à quoi correspond cette mesure lorsqu'il s'agit d'écrans tactiles multi-points. En effet, en considérant dix points de contact, correspondant aux doigts des deux mains, et supposant deux dimensions par point (x et y), le nombre de degrés mesurés par le périphérique utilisateur correspondrait à 20. L'agrandissement d'un objet graphique (largeur et hauteur) par un geste d'écartement, impliquerait un degré d'intégration de $2/20$. Ceci est sans compter que la partie logique de l'instrument tactile, les algorithmes de transformation, reposent d'une part sur la valeur des positions de point en entrée mais également sur l'interprétation de la combinaison de ces points (« dynamicité » interne de la main).

En effet, les quelques 23 degrés de liberté de la main et la bi-manualité de certaines interfaces, autorisent des milliers de configurations et de combinaisons différentes. Le degré d'intégration extrêmement faible de l'instrument tactile crée inévitablement les conditions d'une exploration plus difficile de l'interface. D'une part, sans aide contextuelle il devient plus complexe pour l'utilisateur de comprendre l'association implicite entre l'instrument logique et l'instrument physique et d'en déduire le geste à effectuer. D'autre part, de

multiples combinaisons (une main ou deux main, deux doigts ou plus) concourent à la réalisation d'une même tâche et cela diminue la capacité d'apprentissage de l'utilisateur puisque les liens de cause à effet perdent de leur visibilité.

3.5. Le degré de compatibilité

Le degré de compatibilité¹⁷⁸ mesure la similitude entre les actions physiques effectuées sur l'instrument et le comportement de l'objet d'intérêt. D'un autre côté, l'utilisation d'une barre de défilement verticale implique un faible degré de compatibilité dans le sens où lorsque l'on déplace la barre vers le bas, la fenêtre défile vers le haut. Dans le cas des interfaces tactiles, sélectionner un objet graphique et le déplacer à l'écran implique un fort degré de compatibilité car l'objet doit avoir les mêmes caractéristiques de déplacement que l'instrument d'interaction : position, vitesse, accélération.

Je me permets ici de relater une anecdote familiale car elle illustre parfaitement cette notion. C'était vraiment la toute première fois que je demandais à ma mère d'utiliser un ordinateur. L'objectif, d'apparence triviale, était de pointer une icône située dans un menu en haut de l'interface et de cliquer dessus. Ce qui semblait acquis, voire « naturel » pour moi ne l'était certainement pas pour un novice. En effet, au lieu de pousser la souris vers l'avant sur le plan horizontal du bureau, elle a littéralement soulevé la souris dans les airs et l'a déplacée sur son axe vertical, pensant tout naturellement que le curseur graphique de la souris allait en faire autant. Comment pourrais-je le lui reprocher aujourd'hui, sachant que son geste était plein de bon sens et qu'il faisait figure d'avant-garde si l'on considère les travaux en IHM aujourd'hui. Sans le savoir, ma mère avait imaginé la première souris gyroscopique.

Dans les interfaces WIMP classiques, déplacer une souris sur un plan horizontal (le bureau physique) modifie la position d'un curseur sur un plan vertical (l'écran). Cette indirection spatiale ne nous semble pas forcément gênante car nous nous sommes adaptés à ce fonctionnement, nous en avons fait l'apprentissage. Comme le suppose Anees Ansari¹⁷⁹, le

¹⁷⁸ *ibid.*

¹⁷⁹ ANSARI Anees. *Direct 3D Interaction Using A 2D Locator Device*. Thèse sous la direction de LES Piegl. Department of Computer Science And Engineering, College of Engineering, University of South Florida, 2003

degré d'indirection est largement associé à la charge cognitive nécessaire à l'expérimentation et à l'apprentissage d'un système interactif.

4. Modalités des interactions

4.1. Interaction directe

Contrairement aux interactions indirectes qui consistent à passer par un instrument physique ou virtuel, l'interaction directe suppose de réduire l'écart perceptuel entre le geste de l'utilisateur et le résultat visuel ou sonore opéré sur le système. Les gestes de manipulation des objets graphiques en interaction tactile sont caractérisés par l'absence de périphérie physique indirecte en ce sens où les capteurs sont transparents pour l'utilisateur. Celui-ci agit directement sur l'objet d'intérêt selon une première phase d'approche et de ciblage naturelle par le même geste de préparation qui précède généralement la saisie d'un objet physique. Les interactions se réalisent alors en contact direct avec l'objet à manipuler et le retour, ou feedback tactile et visuel, donne à l'utilisateur la sensation d'un contrôle direct sur ses caractéristiques (forme, position, etc.). Les retours synchrones du système permettent à l'utilisateur de ressentir de manière immédiate les effets de l'interaction et d'en prendre ainsi toute la dimension, par tâtonnement ou par intuition. D'une part, le faible degré d'indirection des systèmes tactiles est une caractéristique essentielle pour créer des interactions plus naturelles. D'autre part, l'absence de périphérie intermédiaire réduit l'écart perceptivo-moteur, c'est-à-dire la différence morphologique entre l'action physique réalisée sur le périphérie d'entrée et le résultat visible, ou sensible à l'écran.

4.2. Morphologie et usage

Les avancées techniques de ces dernières années en matière d'interface tactile et l'adoption progressive de cette technologie par les utilisateurs ont conduit les industriels à intégrer, voire à remplacer les périphériques d'entrées traditionnels (claviers, souris, boutons physiques), dans les dispositifs interactifs. Au cours du chapitre II, nous avons vu que les recherches sur les technologies tactiles ont abouti sur des techniques et des configurations

différentes en fonction de contraintes techniques (qualité des données tactiles, efficacité de la détection, robustesse) ou selon la configuration du produit (miniaturisation, portabilité, etc.) En fonction de la morphologie et de l'usage des dispositifs, un type de périphérique d'entrée tactile sera préféré à un autre.

Cette différence pourrait demeurer insignifiante si l'on considérait que l'utilisateur était déjà suffisamment occupé à s'approprier de nouvelles techniques d'interaction sans qu'on lui impose en plus de saisir leur nature. Toutefois, en considérant les différences morphologiques et techniques des dispositifs tactiles, nous pouvons faire le constat qu'elles suscitent d'une part des modalités d'interactions qui leur sont propres et qu'elles imposent aux designers d'interface de prendre en compte ces différences pour la conception de nouveaux modèles d'interactions. Il serait alors présomptueux de penser qu'une heuristique d'interaction unique serait applicable à l'ensemble des systèmes interactifs multi-tactiles. La configuration d'un dispositif influence considérablement les conditions de son utilisation et dans le cas des technologies tactiles, certains facteurs sont déterminants car ils influencent considérablement les modalités des interactions gestuelles. Il s'agit de propriétés intrinsèques au dispositif comme la taille de la surface interactive, la technologie tactile embarquée, la portabilité du dispositif ou encore son inclinaison par rapport à l'utilisateur. Je propose de dresser une typologie des interactions possibles selon divers dispositifs tactiles.

La taille de la surface interactive peut varier de quelques pouces à plus d'une centaine en fonction du dispositif. Les plus petits d'entre eux, à l'instar de l'*iPod nano*, ne mesurant seulement qu'un pouce et demi, ne disposent pas d'une taille suffisante pour permettre des interactions à plus de deux doigts simultanément. De plus, l'amplitude des mouvements, relativement restreinte, diminue la précision d'une interaction dynamique puisque l'interprétation des mouvements se fait sur une échelle très faible. Dans ce cas, les mouvements du bras sont quasi-inexistants puisque les degrés de liberté des doigts suffisent à eux seuls pour couvrir tout l'écran. Les petites surfaces ont également la particularité de n'afficher que de petits objets graphiques qui peuvent dans le cadre d'une interaction tactile, être recouverts par le doigt lors de la manipulation. Le manque de visibilité des actions de l'utilisateur sur l'objet de la manipulation entraîne une perte significative du feedback visuel dont le rôle est pourtant si précieux en tant que retour immédiat des actions effectuées par l'utilisateur.

Les écrans des lecteurs de médias ou des téléphones portables, dont la taille varie généralement entre trois et quatre pouces, autorisent davantage de liberté de mouvements. Le nombre de points de contact que la technologie utilisée est capable de détecter simultanément est généralement lié à la taille. L'écran de l'*iPhone 3GS* permet par exemple de distinguer jusqu'à cinq points de contact différents. Dans ce cas, la quasi-totalité des interactions se fera avec une main et plus précisément avec un ou deux doigts puisque l'anatomie de la main offre davantage de liberté aux mouvements grâce à l'opposition du pouce et de l'index, conditions de la préhension chez l'homme. La combinaison de ces deux doigts est particulièrement intéressante pour la manipulation d'objets graphiques comme le pincement (ou *pinch*) pour zoomer, ou la rotation du poignet pour effectuer une opération de rotation selon deux points de contacts. Pour les interactions à un doigt, l'index est celui qui est généralement utilisé par l'utilisateur, étant donné que c'est le doigt de la main dont l'homme se sert par défaut afin de montrer une direction ou pour appuyer sur un bouton. Dans le cas des téléphones mobiles, la portabilité de ces dispositifs permet de les tourner à 90° afin de les tenir à deux mains. Dans ce cas, il est courant de voir des interactions réalisées avec les pouces. C'est le cas de certaines interfaces de jeux vidéo dont les commandes sont représentées en bas de l'écran sous la forme de joysticks virtuels manipulables avec le pouce.

Les tablettes présentent une taille d'écran plus large, d'environ dix pouces, qui offrent l'avantage de pouvoir utiliser les deux mains de manière simultanée. C'est le cas par exemple de l'*iPad* d'*Apple* qui dispose d'une technologie d'écran capacitif capable de détecter jusqu'à onze points de contact simultanés. Cette fonctionnalité permet notamment d'utiliser un clavier virtuel en disposant ses mains comme dans le cas d'un clavier physique ordinaire. La taille de la surface interactive est également suffisante pour manipuler deux objets graphiques de manière indépendante et autorise l'utilisation d'interactions plus complexes à trois doigts et plus.

Enfin, les écrans larges, à partir de trente pouces, se déclinent principalement sous deux formes en fonction de leur inclinaison : les surfaces horizontales comme les tables interactives et les surfaces verticales comme les écrans ou les murs d'images. D'un côté, la taille importante de ces dispositifs offre à l'utilisateur l'avantage de pouvoir utiliser sans contrainte ses deux mains et de bénéficier d'interactions bi-manuelles. De plus, en fonction du nombre de points détectés par la surface interactive, plusieurs utilisateurs peuvent

interagir simultanément sur des espaces d'interactions individuels. Dans certains cas, il est même possible pour le système de reconnaître quel est l'utilisateur à l'origine d'un point de contact (cf. *Mitsubishi Diamond Touch*). D'un autre côté, la taille excessive d'une surface interactive peut entraîner certains problèmes d'ergonomie. Comme le font remarquer Shahzad Malik, Abhishek Ranjan et Ravin Balakrishnan¹⁸⁰, certaines situations peuvent nuire à l'interactivité comme atteindre des cibles distantes, naviguer d'un côté à l'autre de l'écran ou tout simplement avoir une vue d'ensemble de l'interface qui peut être rendue difficile si la taille d'écran est trop large et si l'interface n'est pas suffisamment fragmentée.

Les interactions sur de grandes surfaces impliquent des mouvements de l'ensemble du corps : se rapprocher de l'écran, s'en éloigner, aller vers la gauche, aller vers la droite, lever et tendre le bras pour saisir un objet graphique trop éloigné, se baisser pour saisir une icône trop basse, sont autant de mouvements qui diminuent les performances de l'utilisateur et augmentent rapidement le stress et la fatigue musculaire de l'utilisateur. Il est possible dans le cas des tables interactives de s'asseoir autour de la table et de réduire ainsi les coûts de fatigue physique générée par les interactions. Cependant, la position statique de l'utilisateur deviendra une contrainte lorsque celui-ci devra manipuler un objet graphique hors de portée ou si celui-ci produit des points de contacts involontaires en s'appuyant sur la surface ou par le contact de la manche d'une veste par exemple.

Détailler les interactions selon la morphologie des dispositifs nous a permis de mettre en exergue plusieurs propriétés qui sont autant de critères nécessaires à la classification des technologies multi-tactiles. Jusqu'ici, nous avons identifié les caractéristiques suivantes :

- La taille de la surface interactive,
- La portabilité du dispositif,
- L'inclinaison de la surface,
- La manualité (doigts, main dominante / non dominante),
- La bi-manualité (deux mains sont utilisables simultanément),
- La place pour plusieurs utilisateur.

¹⁸⁰ MALIK Shahzad, RANJAN Abhishek et BALAKRISHNAN Ravin. *Interacting with large displays from a distance with vision-tracked multi-finger gestural input*. In *Proceedings of the 18th annual ACM symposium on User interface software and technology (UIST '05)*. ACM, New York, USA, 2005, pp. 43-52

Chaque dispositif offre des caractéristiques et des technologies que d'autres ne possèdent pas forcément en fonction des usages qu'ils impliquent. Dans l'idée de développer une application multi-tactile, il serait dommage de ne pas pouvoir tirer profit de la totalité des avantages et des fonctionnalités de la technologie utilisée. Les différences fondamentales relevées précédemment nous invitent à penser qu'un développement multiplateforme n'est pas concevable et qu'il appartient au designer d'interface d'adapter les modalités d'interaction en fonction du dispositif spécifique qui a été choisi. Comme le soulignait déjà Baeker en 1980, « *bien que la portabilité soit facilitée par l'indépendance du développement vis-à-vis des terminaux, l'interactivité et l'usabilité sont améliorées par un développement spécifique au terminal.*¹⁸¹ » Il s'agit principalement de pouvoir s'abstraire des standards gestuels afin de profiter des spécificités propres aux dispositifs et voir ainsi émerger de nouveaux modèles gestuels et interactifs.

¹⁸¹ « *Although portability is facilitated by device-independence, interactivity and usability are enhanced by device dependence.* » In BAECKER R. M. *Towards a characterization of graphical interaction*. In *Human-computer interaction*. Morgan Kaufmann Publishers Inc., San Francisco, USA, 1987, pp. 471-482



**Lecteur média portable
Apple iPod nano - 1.5"**



**Souris tactile
Apple Magic Mouse**



**Téléphone portable
Apple iPhone - 3,5"**



**Pavé tactile - Apple Magic Pad
13,01x13,13cm**



**Tablette tactile - Motorola Xoom
10.1"**



**Ordinateur tout-en-un
HP Touchsmart - 23"**



**Table interactive - Microsoft Surface
30 à 40"**



Ecran multi-tactile LG - 46"



**«mur» d'image multi-tactile
image > 100"**

Figure 55 - Éventail de dispositifs multi-tactiles. De la plus petite surface à la plus grande. Surfaces horizontales et verticales.

4.3. Définition de données

Les technologies tactiles diffèrent par leur configuration et par la méthode de détection des points de contact. Pendant longtemps, la plupart n'ont été capables de distinguer que les événements de contact et leur position. Lorsqu'un point de contact est détecté, modifié dans l'espace ou supprimé, les capteurs renvoient l'événement au système en l'associant à un numéro d'identification de point unique. Cependant, selon le type de capteurs, le périphérique tactile peut fournir au système davantage d'informations. Il est nécessaire d'identifier préalablement les caractéristiques d'une technologie tactile afin de connaître ses contraintes et le degré d'information du geste qu'elle est capable de révéler. En exploitant la totalité des informations issues des capteurs, il est possible d'enrichir l'expérience utilisateur.

Dans l'idée de concevoir un intergiciel multi-tactile à partir des éléments issus de ce modèle d'interaction, je propose de faire l'inventaire des types de données qu'il est possible d'exploiter pour le designer d'interaction :

- *La position du point de contact.* Selon le type de dispositif, la position du point de contact est plus ou moins précise. Dans le cas des systèmes optiques par caméra, la précision dépend de la qualité de la calibration du système et de la résolution de capture de la caméra. Dans le cas des cadres optiques, la position est déterminée par triangulation des données issues des capteurs. Tout comme la majorité des technologies utilisant un plan de lumière infrarouge couvrant la surface, la détection des positions peut-être perturbée par deux facteurs. Dans le cas des cadres optiques infrarouges, la méthode de détection peut entraîner ce qu'on appelle des points fantômes (*ghost touches*) qui sont présents lorsque deux doigts sont posés au même moment (délai inférieur à 20ms en moyenne). Ce problème peut être corrigé de manière logicielle ou tout simplement en ajoutant des capteurs supplémentaires. Ensuite, selon le type de configuration, il peut y avoir des problèmes d'occlusion, c'est-à-dire qu'un doigt peut en masquer un autre et le microcontrôleur n'a pas suffisamment de données pour effectuer la triangulation.
- *La pression du point de contact.* En connaissant la pression exercée de manière individuelle et dans le temps sur un point particulier de la surface, il est possible

d'offrir des interactions plus riches à l'utilisateur. D'une part, en connaissant la pression d'un point de contact, il est possible d'ajouter des interactions parallèles comme modifier l'épaisseur d'un trait dans une application de dessin ou d'écriture. D'autre part, en couplant la pression des points de contact à un effet dynamique ou visuel sur l'objet du contact, cela augmente considérablement le feedback visuel et donne à l'utilisateur un sentiment de contrôle plus important. Toutefois, la quasi-totalité des surfaces interactives ne détectent pas vraiment la pression réelle appliquée mais en font l'estimation en fonction de la taille de la zone de contact. En effet, les surfaces sont généralement rigides afin de prévenir tout dommage sur l'écran et même si l'utilisateur bénéficie d'un feedback visuel, le feedback tactile reste très limité. Bill Buxton¹⁸² préfère d'ailleurs utiliser le terme de « degrés de contact » (*degree of touch*) plutôt que celle de pression, notant la différence de feedback tactile entre ces deux notions.

- *La taille de la zone de contact.* Cette information est utile pour déterminer approximativement la pression exercée par l'utilisateur sur le point de contact si le périphérique ne fournit pas cette information. Il peut être utile également de connaître cette zone afin de différencier un doigt d'un objet posé sur la surface. Selon le cas, il est possible de fournir un feedback visuel pour informer l'utilisateur que le système a bien reconnu un objet en entourant la zone par un indicateur graphique approprié.
- *L'angle d'approche des doigts.* Certains systèmes présentent l'avantage de pouvoir déterminer l'angle d'approche du doigt, voire de la main de l'utilisateur. Cette donnée à l'avantage de permettre des interactions plus riches avec seulement un point de contact comme par exemple une rotation sur un point unique. Cette information permet également de pouvoir classer des points de contact selon la main qui en est à l'origine. En effet, en connaissant les degrés de liberté des doigts de la main et en comparant l'angle d'approche de chaque point, il est possible de savoir s'ils appartiennent à une même main.

¹⁸² BUXTON Bill. *Multi-Touch Systems that I Have Known and Loved*. Microsoft Research, 2007, 16p. disponible ici : http://www.ekatetra.com/downloads/Chorded_History_BuxtonMultiTouch.pdf

- *Le vecteur de déplacement.* Certains systèmes donnent des informations supplémentaires sur le déplacement d'un point de contact dans le temps comme sa direction, sa vitesse et son accélération. Bien que ces données soient particulièrement utiles, il est toujours possible de les calculer a posteriori côté application si le dispositif ne fournit pas ces informations. Il s'agit simplement de déterminer le vecteur de déplacement du point de contact en fonction de sa position dans le temps.

5. Taxinomie du geste

5.1. Formalisation syntaxique

La richesse des interactions gestuelles permet aux concepteurs d'interfaces de proposer de nouvelles modalités pour interagir avec le système, au-delà des interactions plus répandues comme pointer, déplacer, tourner et redimensionner des objets graphiques. Cependant, la démarche du développeur doit s'inscrire dans une logique d'abstraction où les composantes spatiales et temporelles du geste ne sont disponibles que sous la forme de données brutes (*raw data*) qui sont analysées et interprétées par les algorithmes informatiques. Afin d'aider les développeurs à traiter tous les gestes potentiellement permis par le système, Dietrich Kammer *et al.*¹⁸³ proposent une formalisation des interactions multi-tactiles pour la conception d'environnements de développement (*framework*) à base de briques gestuelles.

Afin de formaliser la syntaxe des interactions multi-tactiles gestuelles, les auteurs s'inspirent de la taxinomie proposée par Wobbrock¹⁸⁴ en décomposant l'interaction par point de contact et par nature (discrète ou continue). La posture correspond au type d'objet en contact avec la surface, il peut s'agir d'un doigt, de la main ou tout autre objet dont la forme est communément appelé *blob*.

¹⁸³ KAMMER Dietrich, WOJDZIAK Jan, KECK Mandy, GROH Rainer et TARANKO Severin. *Towards a formalization of multi-touch gestures*. In *ACM International Conference on Interactive Tabletops and Surfaces (ITS '10)*. ACM, New York, USA, 2010, pp. 49-58

¹⁸⁴ Wobbrock Jacob O., Morris Meredith Ringel et Wilson Andrew D. *User-defined gestures for surface computing*. In *Proceedings of the 27th international conference on Human factors in computing systems (CHI '09)*. ACM, New York, USA, 2009, pp. 1083-1092

ELEMENT DE LANGAGE	INTENTION	EXEMPLE
Posture	Décrit la forme de la zone de contact qui est analysée	2D(...), 1M(....)
Geste atomique	Décrit le mouvement du blob	Déplacement, tenue, ligne, cercle, demi-cercle
Opérateur de composition	Décrit la progression temporelle du geste (parallèle, successif ou asynchrone)	Virgule, astérisque, signe plus
Centre d'intérêt	Spécifie le centre d'intérêt du geste (objet, un groupe d'objet ou une vue utilisateur)	O, G, V
Contrainte de zone	Décrit le mouvement relatif des gestes atomiques (superposé ou convergent)	CROISE[...], SYNCH[...], JOINT[...], ECARTÉ[...]

Tableau 2 - Élément de langage de la formalisation de Kammer *et al.* In [KAMMER, 2010].

Dans la formalisation de Kammer *et al.*, les mouvements sont décomposés sous la forme de gestes atomiques, centrés sur une rose des vents. Il peut s'agir de lignes, de cercles ou demi-cercles qui peuvent être combinés pour créer des phrases gestuelles plus complexes. Les opérateurs de composition permettent de représenter des contacts simultanés et leur gestuelle respective. Le signe multiplié (ou astérisque) correspond à un mouvement synchrone, le signe plus à un mouvement asynchrone. Si les mouvements ne sont pas parallèles, la virgule sert de délimitation.

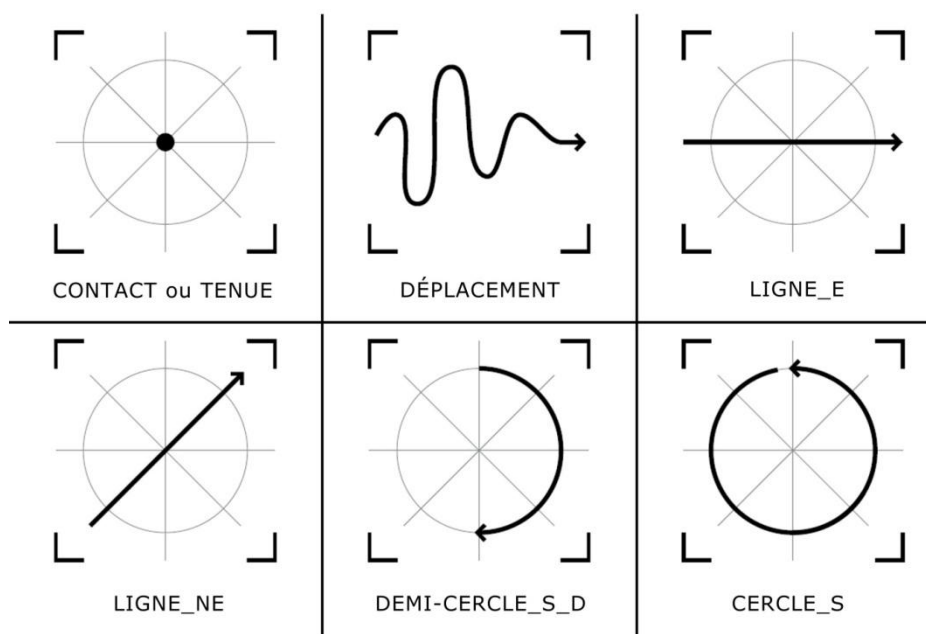


Figure 56 - Exemple de geste atomique. Reproduction. In [KAMMER, 2010].

La réduction de phrases gestuelles complexes en entités atomiques permet de couvrir une quantité importante d'interactions existantes ou potentielles. Toutefois, il est possible que certains gestes trop complexes ne puissent être représentés grâce à cette syntaxe, notamment si sa structure temporelle n'est pas linéaire ou si le type de mouvement est trop spécifique. En effet, cette formalisation ne rend pas compte de la granularité du geste, comme la dynamique du mouvement (vitesse, accélération) ou les formes spatiales plus complexes du mouvement. De plus, cette formalisation ne s'applique qu'aux interfaces pour un utilisateur unique et il serait nécessaire de la compléter afin qu'elle puisse s'étendre aux interactions multi-utilisateurs. Toutefois, cette formalisation présente un intérêt non négligeable lors de la conception et de l'écriture des phrases gestuelles par le designer d'interactions en ce sens où elle permet de décomposer les traits atomiques du geste et de simplifier le développement d'algorithmes d'interprétation du mouvement.

5.2. Taxinomie des interactions multi-tactiles

Précédemment, nous avons écarté de notre définition le caractère intuitif du geste en rapprochant davantage le terme « naturel » à celui qui profite des compétences préalablement acquises par l'homme au cours de son existence. Jef Raskin tranche sur le sujet en déclarant que l'homme ne possède pas de telles aptitudes, qu'il ne peut « [...] *acquérir de la connaissance sans exposition préalable du concept, sans devoir passer par un stade d'apprentissage, et sans avoir à réfléchir de manière rationnelle*¹⁸⁵ » Il s'agit de mettre de côté la définition philosophique de l'intuition comme « *connaissance immédiate qui ne recourt pas au raisonnement*¹⁸⁶ » et comprendre cette notion comme le pressentiment de ce qui est ou doit être. En ergonomie, l'intuitivité reflète donc davantage le caractère de prédictibilité de l'action des interactions de l'utilisateur sur les actions du système.

Je souhaitais dans ce chapitre pouvoir classer les gestes tactiles selon leur degré de simplicité. Toutefois, la nature protéiforme du geste en interaction homme-machine est soumise à tant de facteurs hétéroclites qu'il apparaîtrait rapidement autant de classes qu'il

¹⁸⁵ « *However, there is no human faculty of intuition, as the word is ordinarily meant; that is, knowledge acquired without prior exposure to the concept, without having to go through a learning process, and without having to use rational thought.* » In RASKIN Jef. *The Humane Interface : New directions for designing interactive systems*. Addison-Wesley Professional, 2000, 256 p.

¹⁸⁶ Dictionnaire Historique de la Langue Française. Dictionnaire historique. Sous la direction d'Alain Rey. Le Robert, 2006, 4304 p.

existe de gestes différents. En effet, la complexité d'un geste est intimement liée aux facultés cognitives de l'homme, différentes d'un individu à un autre, à savoir sa capacité de mémorisation, d'adaptation et d'apprentissage. D'autre part, elle est tout autant déterminée par les aptitudes physiques de l'utilisateur comme l'habileté, l'adresse et la précision nécessaire à sa réalisation. Enfin, la complexité du geste dépend de sa forme spatiale et temporelle car plus le geste contient de mouvements (*strokes*), plus sa facilité d'apprentissage est réduite.

Il existe également un phénomène de standardisation de certains gestes qui sont le résultat de la popularité de systèmes tactiles comme l'*iOs d'Apple* et qui peuvent nuire à la classification. Dans ce cas précis, la diffusion très large des téléphones de la marque a fortement contribué à l'adoption d'un style d'interaction qui tend à se généraliser. Un exemple probant est le geste de pincement (ou *pinch*), qui consiste à rapprocher ou écarter le pouce et l'index simultanément afin de modifier la taille d'une image ou de la vue utilisateur. Nous serions tentés d'ajouter, comme facteur de transparence, l'appartenance d'un geste à un standard. Cependant, l'idée de définir un modèle d'interaction multi-tactile représente pour moi non seulement la volonté de créer un espace de compréhension des systèmes existants mais également d'être en mesure de créer de nouveaux styles d'interaction. Dans le cadre de l'approche conceptuelle que je défends, je ne tiendrai pas compte des styles d'interactions dominants afin de conserver la plus grande latitude possible pour la phase de développement. Ce n'est qu'après l'expérimentation et l'extension de notre modèle qu'il nous sera possible de mesurer l'influence des styles d'interaction dominants comme celui de l'*iPhone*.

La standardisation des interactions est d'une manière ou d'une autre inévitable et c'est en partie nécessaire pour ne pas créer la confusion chez les utilisateurs qui se retrouveraient confrontés à des modalités d'interaction différentes d'un système à un autre. Toutefois, une standardisation trop précoce pourrait bloquer la technologie à un stade primitif où certaines règles d'interaction pourraient se révéler inefficaces avec le temps, comme cela a été le cas pour la répartition des lettres sur nos claviers dont l'origine remonte aux contraintes mécaniques des machines à écrire. Car une fois qu'un modèle s'est constitué comme standard, il devient plus long et plus difficile de les faire évoluer.

Suite à ces considérations, il me semble plus logique de classer les gestes, non pas selon leur simplicité, toute subjective, mais en m'inspirant de la classification de George et Blake qui distinguent les gestes selon leur nature fonctionnelle, détaillée précédemment dans la classification de Claude Cadoz, à savoir les gestes de la manipulation et les gestes sémiotiques (gestuelle). A cette distinction, nous ajouterons deux catégories qui nous semblent relever d'une nature légèrement différente : les gestes d'écriture et ceux mimant les interactions WIMP traditionnelles.

Gestes traditionnelles :

- *Le geste d'écriture.* Il s'agit du mode linguistique qui permet d'écrire directement sur l'écran afin que le système procède à une reconnaissance de caractères. La difficulté de ce mode réside davantage côté système dans le sens où les algorithmes peinent parfois à interpréter l'écriture naturelle, le cursif lié, obligeant parfois l'utilisateur à écrire des caractères séparés. La contrainte pour l'utilisateur réside dans la nécessité bien souvent d'entraîner le système à reconnaître son écriture avec plus de précision. Bien que la reconnaissance de caractères existe dans les interfaces traditionnelles (tablette graphique, assistant personnel, etc.), les technologies multi-tactiles où le « geste nu » est sollicité, sans stylet ou accessoire, l'écriture avec le doigt peut vite se révéler imprécis et bien moins naturel.

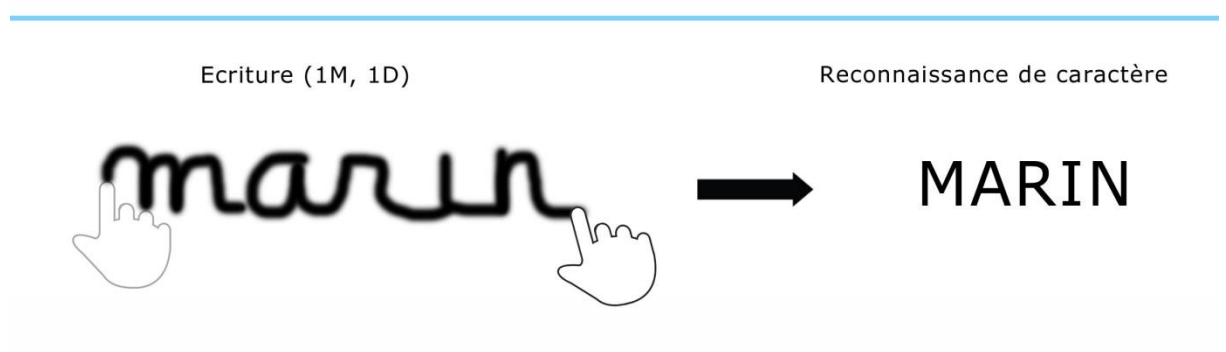


Figure 57 - Illustration du geste d'écriture.

- Le geste de dessin. C'est principalement le geste qui permet de dessiner sur la surface interactive. Les écrans multi-tactiles offrent la possibilité de dessiner avec chaque doigt de la main, voire avec les deux mains mais présentent le désavantage que la taille du doigt n'offre pas la même finesse de dessin que le stylet d'une tablette graphique par exemple. Toutefois, l'interface multi-points peut donner lieu à de nouvelles formes d'interactions comme modifier les propriétés de trait et la couleur du pinceau avec une main pendant que l'autre dessine. Il est également possible de reconnaître la forme des figures dessinées par l'utilisateur afin d'en créer des commandes système.

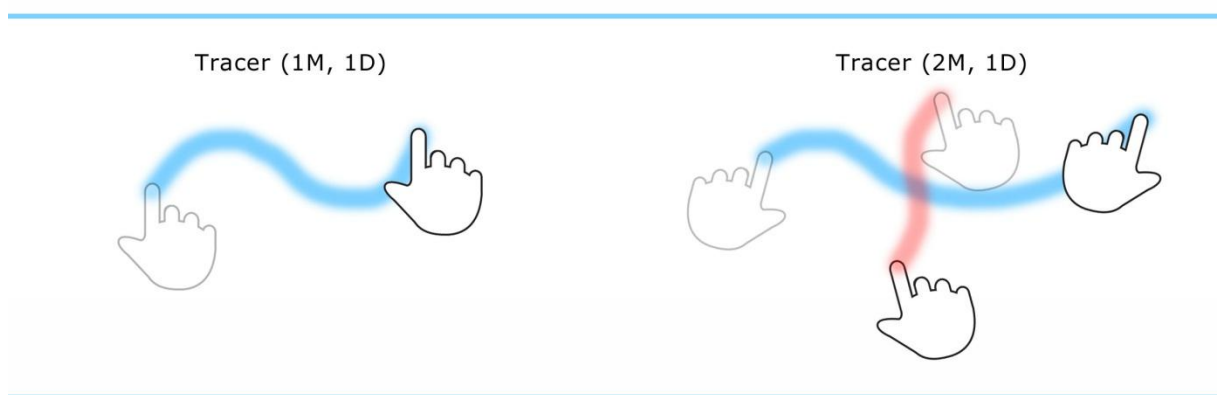


Figure 58 - Illustration du geste de dessin. Manuel et bi-manuel.

Geste de manipulation :

- *Interactions de style WIMP*. Les gestes de manipulation concernent ceux qui permettent de modifier les propriétés d'éléments graphiques en mettant les doigts ou la main directement en contact avec l'objet d'intérêt. Parmi ces gestes, certains se distinguent par leur similitude avec des interactions classiques à la souris, comme cliquer sur un bouton, appuyer et relâcher la souris ou réaliser un glisser / déposer. La différence réside tout d'abord en ce que l'interaction tactile gestuelle propose une interaction directe, spatialement et temporellement, puisque l'objet d'intérêt est situé juste sous les doigts de la main. Ensuite, ces interactions peuvent être effectuées avec plusieurs points de contact, ce qui suppose la possibilité de créer des interactions spécifiques en fonction du nombre de doigts posés sur l'objet à manipuler. Par exemple, la désignation d'un objet se fait à un ou plusieurs doigts et

le glisser / déposer n'est réalisé qu'à partir du moment où deux doigts ou plus sont posés à l'écran.

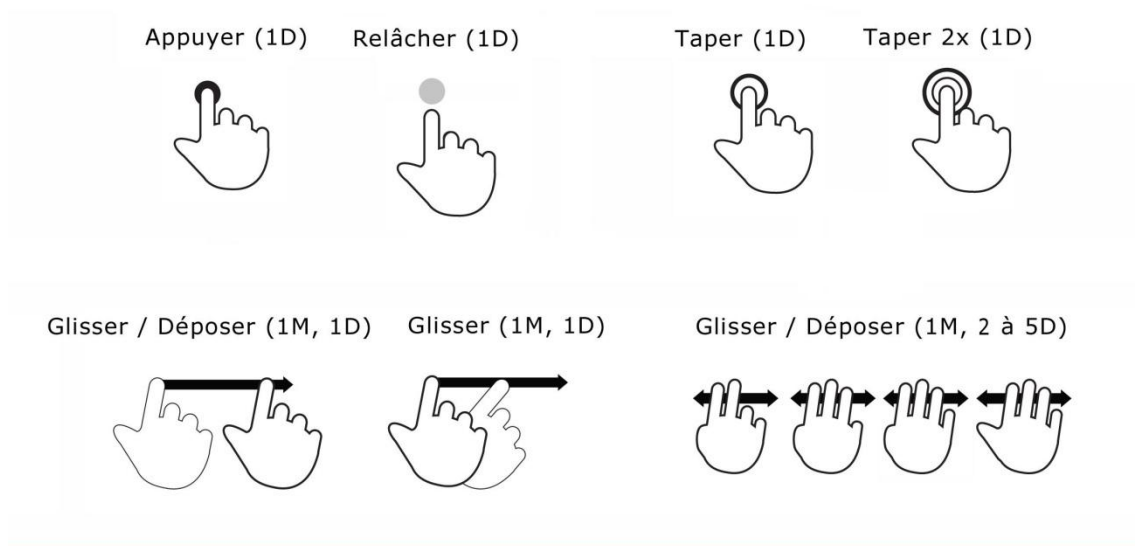


Figure 59 - Illustration de quelques gestes de manipulation rappelant le style WIMP.

- *Interactions de manipulation multi-tactiles spécifiques.* D'autres gestes de manipulation, au contraire, ne correspondent à aucune interaction traditionnelle étant donné qu'ils sont réalisés grâce à plusieurs points de contact ou par l'apposition des mains sur l'écran. De plus, il est possible de caractériser ces gestes selon leur couplage avec le système. Selon le scénario d'interaction, un geste peut être associé à une ou plusieurs tâches simultanées. Plusieurs gestes peuvent quant à eux être associés à une tâche unique ou à plusieurs tâches distinctes.

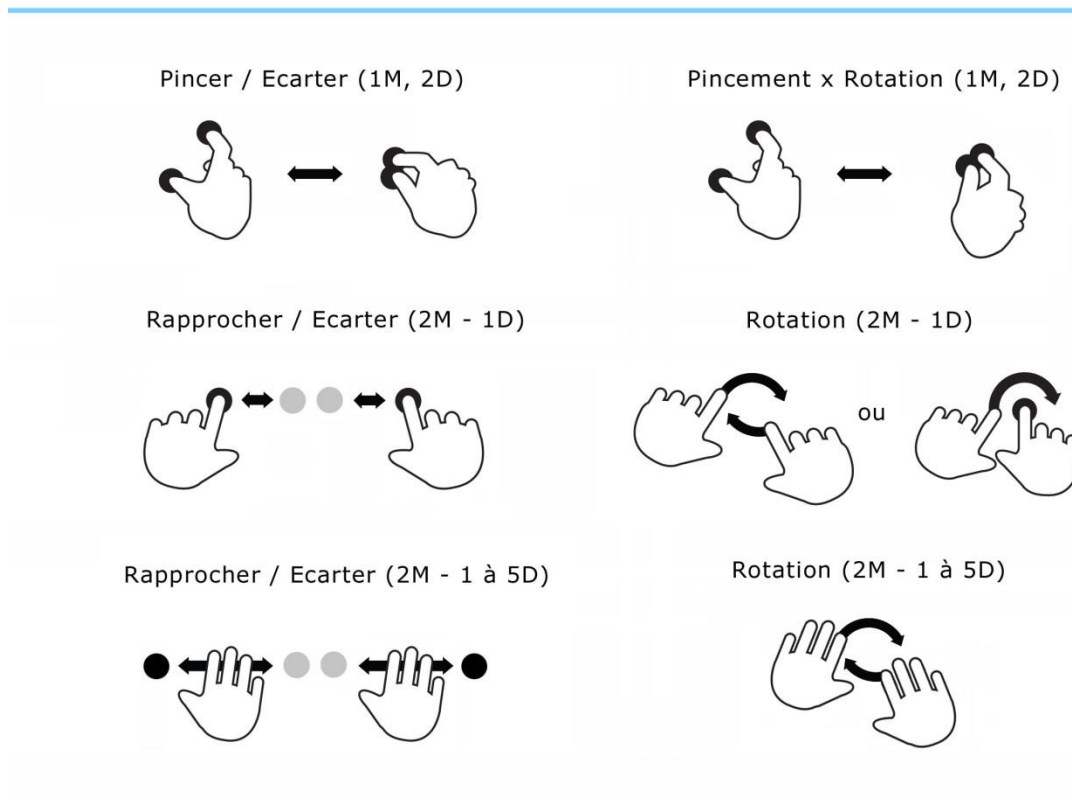


Figure 60 - Illustration de quelques gestes de manipulation spécifiques.

Gestes sémiotiques

Jacob O. Wobbrock *et al.*¹⁸⁷ ont proposé une classification des gestes sémiotiques qui distinguent trois catégories principales. D'une part, les gestes symboliques qui tracent à l'écran l'icone associée à une commande. Par exemple, tracer une flèche vers la gauche pour effectuer une commande de retour dans la navigation ou dessiner une croix à l'écran pour supprimer un objet ou annuler une opération en cours. D'autre part, les gestes métaphoriques qui interviennent lorsqu'ils simulent l'action sur, avec, ou comme un autre objet. Par exemple, tracer un cercle en continu peut représenter l'action sur un anneau de commande circulaire, effectuer un mouvement rapide de la droite vers la gauche peut représenter l'action de tourner la page d'un livre. Enfin, les gestes abstraits qui n'ont pas de dimensions symboliques ni métaphoriques. Ils sont associés de manière arbitraire aux

¹⁸⁷ Wobbrock Jacob O., Morris Meredith Ringel et Wilson Andrew D. *User-defined gestures for surface computing*. In *Proceedings of the 27th international conference on Human factors in computing systems (CHI '09)*. ACM, New York, USA, 2009, pp. 1083-1092

commandes systèmes et possèdent un degré de conventionnalité plus important. Tenir l'index appuyé et taper avec le majeur correspond dans certains systèmes multi-tactiles à l'action du clic droit de la souris.

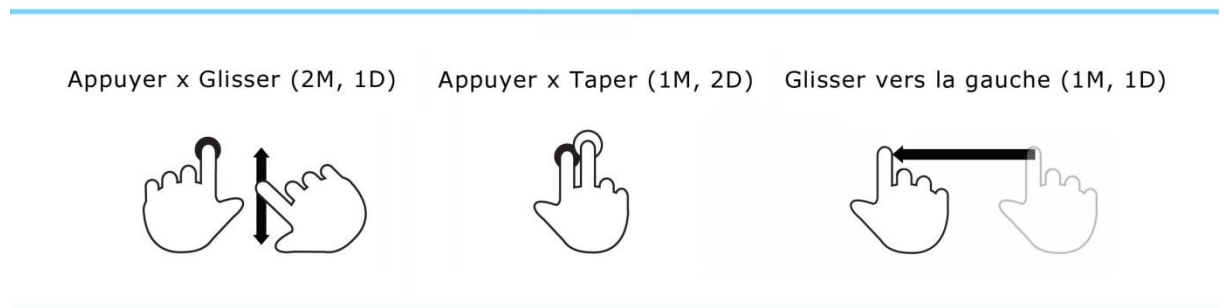


Figure 61 - Illustration de quelques gestes lexicaux.

6. Design d'interaction

L'informatique revêt déjà les caractères de l'informatique ubiquitaire, ou ambiante, en ce sens où les ordinateurs ont envahi nos foyers, nos écoles, nos bureaux et nous accompagnent lors de nos déplacements sous la forme de dispositifs mobiles. L'utilisateur est contraint de se familiariser avec la morphologie des périphériques d'entrée, avec les nombreuses interfaces et méthodes d'interaction propres aux dispositifs ou aux choix des concepteurs.

Dans le contexte du design de systèmes tactiles, il est nécessaire de considérer le style d'interaction, l'environnement graphique, les outils et le couplage entre les interactions utilisateur et les réponses système. Bien souvent, il existe un réel décalage entre le concept d'interface naturelle et la simplicité d'utilisation des applications. Donald A. Norman, professeur au MIT, évoque le sentiment de culpabilité ressenti par les utilisateurs lorsque ceux-ci utilisent des objets ou des appareils du quotidien et se retrouvent dans une situation où ils ne répondent pas correctement aux intentions initiales de leur manipulation. L'utilisateur, sans véritablement l'avouer, se sent coupable des erreurs qu'il commet, des problèmes rencontrés lors de l'utilisation ou du fait de ne pas savoir comment s'y prendre.

Lors des différents salons auxquels nous avons participé, avec l'entreprise partenaire du CIFRE, nous avons fait l'observation, lors des phases préliminaires du projet, que les utilisateurs étaient particulièrement intéressés à l'idée de pouvoir tester les interfaces multi-tactiles que nous leur propositions. Parfois, ces utilisateurs se retrouvaient confrontés à un problème dû à la captation de mouvement, au manque d'explication de notre part, ou tout simplement à des comportements du système qui ne correspondaient pas à l'intention initiale de leurs interactions. Leur réaction était immédiate et sans appel, ils écartaient promptement leurs mains de la surface interactive en s'excusant d'avoir fait une erreur ou en se défendant de ne pas « être doué » avec la technologie. Norman défend l'idée que c'est le designer qui est à blâmer car il a conçu un objet qui n'est pas facile à comprendre et qui est trop ouvert aux erreurs de la part de l'utilisateur. Bien qu'il étende son point de vue à une notion plus générale du design d'objet, il formule quatre principes de conception qu'il nous semble pertinent d'adapter au design d'interactions de systèmes tactiles : la visibilité, le mapping naturel, l'affordance et le feedback.

6.1. Visibilité

Le cycle d'action défini par Norman est un modèle psychologique qui décrit les différentes étapes réalisées par l'utilisateur dans une situation d'interaction avec un système informatique. Ce modèle permet de designer ou d'évaluer les qualités d'usabilité d'une interface en mettant les différents processus de réalisation d'une tâche en fonction de la visibilité de l'interface. Le cycle d'action se décline en sept étapes :

- L'élaboration d'un but, d'un objectif.
- La traduction de l'objectif en intention, en actions spécifiques qui permettent d'atteindre l'objectif fixé.
- La traduction de l'intention en combinaisons de commandes plus détaillées – une stratégie de manipulation de l'interface.
- L'exécution du plan.
- La perception de l'état de l'interface.
- L'interprétation de la perception à la lumière des attentes.
- L'évaluation ou la comparaison des résultats par rapports aux intentions ou à l'objectif initial.

Selon le cycle d'action décrit par Norman, il apparaît nécessaire que l'utilisateur possède une représentation mentale particulièrement claire de l'environnement pour formuler un objectif initial, une compréhension suffisante pour déduire les commandes à effectuer en fonction des objets ou des outils visibles et un retour visuel (ou feedback) pour évaluer les résultats de ses actions. La visibilité, telle que la conçoit Norman, correspond à la lisibilité des actions et des rétroactions qui sont potentiellement réalisables pour atteindre l'objectif prévu. Dans le cadre des interfaces traditionnelles, les actions disponibles sont repérables car elles apparaissent sous la forme de boutons, d'étiquettes, de menus déroulants et d'icônes. De plus les modalités d'interaction à la souris sont relativement limitées à son déplacement, aux deux boutons dont les fonctions de pointage et de propriété sont applicables à de nombreux programmes, et à la molette de défilement dont le nom suggère déjà sa fonction principale.

Toutefois, appliquer ce principe dans la conception d'interfaces multi-tactiles semble plus délicat. D'une part, les interactions sont contextualisées car elles dépendent de l'objet graphique, de l'outil ou de la fenêtre auxquelles elles s'appliquent. D'autre part, le degré de conventionnalité élevé de certains gestes peut nuire à la visibilité des actions si ceux-ci ne sont pas répertoriés et accessibles dans une rubrique d'aide par exemple. Enfin, le nombre élevé de gestes permis par le degré de liberté de la main ou par leur combinaison, ainsi que la similarité de certains gestes peut nuire à la qualité de l'évaluation de l'utilisateur puisqu'il est plus difficile pour l'utilisateur d'identifier l'origine d'une erreur, ne sachant pas si elle provient d'un geste mal réalisé ou d'une interprétation erronée du système.

La visibilité fait également référence à la simplicité apparente du système, en ce sens où son degré de complexité ne doit pas compromettre pour autant la simplicité de son utilisation.

« La notion de complexité fait référence au fonctionnement interne du système, la difficulté, à la face visible par l'utilisateur – les facteurs qui affectent la simplicité d'utilisation. L'histoire des technologies démontre que le moyen de créer des usages plus simples, moins difficiles, fait souvent appel à des systèmes plus sophistiqués, intelligents et complexes. Avons-nous besoin d'interfaces intelligentes ? Je ne pense pas : L'intelligence doit rester à l'intérieur, interne au système. L'interface est la partie visible du système, où les utilisateurs ont besoin

de stabilité, de prédictibilité, et de l'image d'un système cohérent qu'ils peuvent comprendre et de ce fait apprendre.¹⁸⁸ »

Les tâches complexes peuvent être simplifiées, restructurées, afin de minimiser la charge cognitive et le temps nécessaire pour trouver la solution de leur réalisation. Norman nous incite à prendre en considération les alternatives proposées par la technologie afin d'aider l'utilisateur à discerner l'implication de ces actions dans le bon déroulement d'une tâche :

- Faciliter l'effort cognitif en fournissant de l'aide.
- Rendre visible ce qui peut améliorer le contrôle utilisateur (feedback).
- Automatiser les tâches.
- Changer la nature de la tâche.

6.2. Mapping naturel

Le concept principal du mapping naturel consiste à ce que la relation spatiale entre la position des contrôles et des objets sur lesquels ils opèrent soit le plus direct possible, avec les contrôles placés soit sur les objets eux-mêmes ou disposés de manière analogue. Dans la même idée, les mouvements des contrôles doivent être analogues ou similaires aux événements attendus du système. Une part critique de l'action est l'évaluation de ses effets, c'est le rôle déterminant du feedback fourni par le système sur les résultats visibles ou audibles des intentions de l'utilisateur. Afin de créer des interfaces plus efficaces, Norman suggère d'exploiter le mapping naturel qu'il définit comme « *profiter de l'avantage des analogies physiques et des standards culturels¹⁸⁹* » pour donner à l'utilisateur une compréhension immédiate du fonctionnement de l'interface.

¹⁸⁸ « *Complexity refers to the internal workings of the system, difficulty to the face provided to the user -- the factors that affect ease of use. The history of technology demonstrates that the way to make simpler, less difficult usage often requires more sophisticated, more intelligent, and more complex insides. Do we need intelligent interfaces? I don't think so: The intelligence should be inside, internal to the system. The interface is the visible part of the system, where people need stability, predictability and a coherent system image that they can understand and thereby learn* ». In NORMAN Donald Arthur. *Complexity versus Difficulty: Where should the Intelligence Be?* . Plenary Address. In *Proceedings of the 7th international conference on Intelligent user interfaces (IUI '02)*. ACM, New York, USA, 2002, pp. 4-4

¹⁸⁹ NORMAN Donald Arthur. *The Design of Everyday Things*. Basic Books, 2002, 272 p.

Exploiter le mapping naturel, c'est s'assurer que l'utilisateur détermine les relations suivantes :

- Entre les intentions et les actions possibles.
- Entre les actions et leurs effets sur le système.
- Entre l'état actuel du système et ce qui est perceptible à la vue, à l'ouïe et au toucher.
- Entre la perception de l'état du système et les besoins, les intentions et les attentes utilisateur.

6.3. Affordance

La notion d'affordance a été introduite par James James Gibson dans l'article « *Theory of Affordances*¹⁹⁰ » et précisée dans son ouvrage sur la théorie de l'écologie de la perception¹⁹¹ qui a considérablement influencé la conception que l'on a de la perception visuelle. Selon lui, percevoir l'environnement conduit à une stratégie de l'action. L'affordance, comme « *indices de l'environnement qui indique les actions possibles* », seraient perçues de manière immédiate, sans interprétation. Inspiré des théories Gestaltistes, l'écologie de la perception place la notion d'affordance comme concept clef de notre activité sensori-motrice.

Dans un essai de 1999¹⁹², Norman apporte un éclairage nouveau sur le terme d'affordance tel qu'il a pu être utilisé dans la littérature propre au domaine des IHM. Il distingue d'une part l'affordance réelle de l'affordance perçue. Dans la théorie de l'écologie, Gibson montre comment une surface plate et rigide permet (*afford*) ou offre à l'animal les moyens d'y reposer. Norman défend l'idée que cette affordance dépend d'une relation purement physique entre l'acteur et l'artefact. Norman distingue la notion d'affordance, qui sans précision, renvoie à l'affordance réelle décrite par Gibson, à l'affordance perçue, qui concerne les caractéristiques de l'apparence d'un dispositif ou d'un objet permettant d'en déduire ses modalités d'action. Le concept d'affordance n'est plus limité aux capacités physiques de l'acteur mais dépend également de facteurs cognitifs comme l'objectif fixé par

¹⁹⁰ GIBSON James J. *The Theory of Affordances*. In *Perceiving, Acting, and Knowing*. Robert Shaw and John Bransford, New York, 1977, pp. 67-82

¹⁹¹ GIBSON J. James. *L'approche écologique de la perception visuelle*. MF, 2010, 332p. (Coll. Dehors)

¹⁹² NORMAN Donald Arthur. *Affordances, conventions, and design*. In *Interactions*, Volume 6 Issue 3, 1999, pp. 38-42

l'acteur, ses pensées et son expérience passée. L'affordance perceptuelle, d'ordre relationnel et subjectif, suggère comment interagir avec un objet.

L'acception du mot proposée par Norman nous donne un cadre théorique pour la compréhension du terme appliqué au domaine des IHM. Par la suite, dans la littérature spécialisée, le terme a été usité pour désigner la propension non quantifiable des caractéristiques visuelles d'un objet à suggérer les actions qu'il permet. Afin d'apporter une lecture plus complète du terme appliqué aux interfaces multi-tactiles gestuelles, et dans la mesure où les interactions impliquées dans ce type de dispositifs concernent des objets virtuels, nous souhaitons mettre en exergue le degré de réalisme, simulé par le système, de l'affordance des objets graphiques.

En effet, le paradigme d'interfaces utilisateur naturelle privilégierait des interactions dont les modalités de manipulation et le comportement des objets graphiques simulent ceux du monde physique. Cette particularité permet de rapprocher l'affordance physique des artefacts virtuels à ceux du réel et permet à l'utilisateur de projeter, dans une certaine mesure, les moyens d'action qui lui sont offerts. C'est le cas par exemple d'une application de visualisation de photos, où celles-ci sont disposées face visible sur une surface, en suggérant à l'utilisateur qu'il est possible de les déplacer sur la surface comme il le ferait avec de vraies photos placées sur une table : du bout du doigt ou avec la main. Cependant, je soulève deux problèmes majeurs liés à l'opposition ontologique entre l'affordance d'un objet réel et celle d'un objet virtuel.

D'une part, l'affordance perçue des objets graphiques, conçus comme des métaphores d'objets du monde réel (bureau, boutons, photos, objets 3D, etc.), et leur comportement, régi par des moteurs physiques (déplacement, inertie, vue en trois dimensions, etc.), créent les conditions d'une simulation réaliste qui donne un cadre de référence à l'utilisateur. Nous suggérons qu'en voulant à tout prix créer des interfaces réalistes, de nouvelles contraintes puissent apparaître et inverser l'ambition initiale de simplification. D'une part, créer un espace virtuel dont l'espace d'activation est régi par les lois physiques du réel revient aussi à intégrer dans l'application informatique les mêmes contraintes qui limitent nos actions avec les objets de notre environnement. Il serait incohérent de ne pas faire profiter les interfaces graphiques du potentiel offert par le numérique et qui n'existe pas dans le réel. Comment en effet peut-on représenter de manière analogue au réel, des actions abstraites comme le

traitement par lot ou les macro-commandes ou des tâches spécifiques aux programmes informatiques, sans déroger au réalisme des interactions. Il est possible qu'en ajoutant des fonctions abstraites, l'utilisateur, sans information supplémentaire, ne fasse pas l'effort cognitif nécessaire à l'exploration de ces fonctions, étant donné qu'il est confortablement installé dans une heuristique d'interactions dont il pense connaître l'étendue.

D'autre part, l'effet inverse peut se produire, étant donné que l'ensemble des phénomènes physiques du réel ne peut exister dans les interfaces virtuelles. La nature d'un dispositif tactile, contrairement à une interface haptique, n'offre pas la même qualité de retour tactile, en termes de texture, de matière, de contour et de forme. De plus, le degré de réalisme graphique et physique des objets virtuels crée l'illusion d'une affordance perçue qui est singulièrement différente de l'« affordance réelle » de l'objet virtuel. L'utilisateur, pourtant conscient qu'il agit sur des objets graphiques représentés sur un écran, éprouve obligatoirement une déception proprioceptive lorsque sa main rencontre une surface plane et sans relief, des objets sans texture et sans volume, alors que sa perception visuelle lui suggère le contraire. Il lui faut alors composer avec des règles de manipulation différentes puisque la physique naïve des comportements interactifs ne suffit pas à créer l'illusion parfaite du réel. L'espace des objets ne se déploie qu'en deux dimensions sur une surface où les objets sont immatériels, insaisissables, et n'offrent aucune prise réelle à la préhension. La perception haptique de l'utilisateur contredit sa perception visuelle et entraîne un « vertige sensitif », résultat de la rupture cognitive de la perception sensori-motrice. La simulation graphique réaliste est trahie par son immatérialité tangible, la simulation graphique se transforme en simulacre proprioceptif.

6.4. Feedback

L'adoption de nouvelles modalités d'interaction par les utilisateurs passe par une phase d'apprentissage qui peut être plus ou moins laborieuse selon le degré de visibilité de l'interface, l'affordance des outils ou des objets à manipuler et dans le cas des interfaces multi-tactiles, le degré de conventionnalité des gestes d'interaction. Ces caractéristiques répondent aux deux premiers critères d'usabilité de l'interface que sont l'efficacité, capacité à atteindre un résultat défini, et l'efficience, capacité à atteindre le résultat avec un effort et un temps minimal. Toutefois, en s'intéressant au critère subjectif de la satisfaction

utilisateur, le reproche couramment répandu, concernant les technologies tactiles, est sa faible propension à fournir un retour tactile comme le feraient par exemple un clavier et une souris traditionnelle. Certains utilisateurs, au-delà de préférer la sensation mécanique de l'action sur un périphérique physique au contact lisse d'une surface tactile, ont l'impression de taper plus rapidement sur un clavier traditionnel que sur un clavier virtuel.

Le feedback, terme introduit en 1914 par E.H. Armstrong et appartenant initialement au domaine de l'électronique, peut se définir comme le réglage des causes par les effets, dans un processus d'autorégulation. Au sens littéral, le feedback signifie « *nourrir en retour* » et désigne l'action en retour d'un effet sur le dispositif qui lui a donné naissance, et donc, ainsi, sur elle-même. Dans le domaine de l'interactivité, nous définirons le feedback comme l'autorégulation des interactions utilisateur par l'action en retour de leurs effets sur le système. La notion de feedback est essentielle pour comprendre comment un utilisateur perçoit les relations de cause à effet, entre ses actions et les réponses du système. Ce retour d'information peut prendre plusieurs formes dont les caractéristiques dépendent autant de la nature du périphérique d'entrée que des réponses système définies par les concepteurs. Dans le cas des interfaces utilisateur tactiles, nous pouvons relever l'influence des retours haptiques, réflexifs, sonores et visuels qui interviennent dans l'usabilité de tels dispositifs.

Feedback haptique.

Le problème que j'évoquais en préambule de cette partie est lié à la faible granularité du feedback haptique, quasi inexistant pour le moment dans les interfaces tactiles. En effet, la surface interactive est plane, sans irrégularité et ne propose pas de retour de force ni de retour sonore comme le fait, de par sa nature mécanique, un clavier traditionnel. Le seul retour qu'elle propose est la sensation de contact avec l'écran qui notifie l'utilisateur qu'il a bien touché un élément graphique de l'interface. Et même dans ce cas, le retour haptique n'est pas toujours synchrone avec le retour visuel car la technologie d'acquisition embarquée dans certains dispositifs ne permet pas de déterminer avec précision le moment exact où le contact est effectué. Il s'agit d'une contrainte technique majeure qui, comme nous l'analysons au chapitre VI, dans le cadre de l'implémentation d'une correction logicielle, peut avoir des conséquences néfastes sur le sentiment de contrôle de la part de l'utilisateur.

D'un autre côté, toujours selon les utilisateurs, la proximité du clavier virtuel par rapport à la zone de texte active, disons son degré d'indirection plus faible, augmente la sensation de contrôle des « fautes de frappe » puisque la vue doit parcourir moins de distance entre chaque vérification d'erreurs. Toutefois, le manque de retour haptique du clavier virtuel peut entraîner un manque de précision dans la réalisation de la tâche et dans la perception des erreurs de frappe. Ce manque de feedback tactile peut facilement se généraliser à l'ensemble des interactions par gestes de manipulation puisque les objets d'intérêt n'offrent aucune réelle résistance physique à la modification de leur position (sensation de friction, de gravité) ou de leur forme (élasticité). De plus le manque de granularité des sensations tactiles va à l'encontre des principes d'interface naturelle puisque les objets graphiques n'offrent aucun retour sur ses qualités intrinsèques comme ses lignes de contour, ses textures ou sa forme spatiale.

Toutefois, il existe des solutions afin de corriger les défauts inhérents aux technologies tactiles, en jouant sur les différentes formes de feedback permis par le dispositif. Norman suggère que « *sans feedback, on se demande sans cesse si quelque chose est en train de se produire ou pas. Peut-être que le bouton n'a pas été poussé assez fort ; peut-être que la machine s'est arrêté de fonctionner ; peut-être réagit-elle de la mauvaise manière.*¹⁹³ » Au cours de nos expérimentations, nous avons mis au jour trois types de feedbacks : réflexif, discret et continu.

Le feedback réflexif

Il s'agit du premier type de notification système qui indique l'état du périphérique d'entrée et ses différentes propriétés. Il reflète d'une certaine manière l'action de l'utilisateur sur le dispositif, en flux continu, même si ces actions ne sont associées à aucune commande particulière. Dans le contexte des interfaces WIMP, la position de la souris est en permanence représentée à l'écran sous la forme d'un curseur. Dans les systèmes d'acquisition gestuelle par caméra, la main est parfois représentée à l'écran par son double numérique et constitue un mécanisme de retour perceptuel qui est nécessaire pour situer le geste dans l'espace d'action.

¹⁹³ « *Without feedback, one is always wondering whether anything has happened. Maybe the button wasn't pushed hard enough; maybe the machine has stopped working; maybe it is doing the wrong thing.* » Traduction libre : Rudy Morin. In NORMAN Donald Arthur. *The Design of Everyday Things*. Basic Books, 2002, 272 p. Traduction libre : COQUELIN Marine et MORIN Rudy.

Dans le cas des interfaces tactiles, la main fait office de pointeur en ce sens où les interactions de désignation et de manipulation sont directes. Le feedback est direct puisqu'il n'existe aucune indirection spatiale entre le périphérique et l'interface. Toutefois, comme nous le relèverons lors des premières observations faites en public avec des utilisateurs non avertis, l'absence de représentation de l'état du système en entrée implique l'absence d'informations contextuelles qui pourraient indiquer à l'utilisateur les modalités d'interaction des objets affichés à l'écran. C'est par exemple le changement d'état du curseur de la souris lorsque celle-ci passe au-dessus d'un bouton, d'un outil ou d'un objet graphique interactif qui avertit l'utilisateur de la nature des actions qu'il est possible d'effectuer (ex : la main avec l'index pointé vers l'avant suggère l'action de cliquer, la main pleine celle d'un déplacement).

Feedback discret

Il s'agit principalement de fournir à l'utilisateur un retour visuel ou sonore lorsqu'une tâche a été exécutée. Ce feedback intervient donc ponctuellement soit en affichant le résultat des actions par la modification des propriétés d'un objet par une interaction discrète (ex : changement de couleur après sélection) soit en indiquant le succès de la tâche prévue par une fenêtre flottante lorsque l'effet de l'action n'est pas visible directement (ex : modification de propriétés internes au système). Enfin, ce retour peut être sonore, c'est le cas notamment des événements sonores qui indiquent à l'utilisateur qu'il vient de taper sur une touche du clavier virtuel. Dans ce cas précis, cette information est précieuse pour l'utilisateur car elle l'informe que la touche a bien été enregistrée. Elle peut également lui signifier rapidement une erreur de frappe si par exemple l'utilisateur tape sur deux touches virtuelles simultanément. D'une certaine manière, le feedback sonore est essentiel pour pallier le manque de retour haptique des dispositifs tactiles. Un clavier ou une souris ordinaire offrent plusieurs niveaux de feedback haptiques. D'une part, les sensations tactilo-kinesthésiques permettent de sentir les contours d'une touche ou d'un bouton, le retour de force permet de savoir à quel moment précis l'action est réalisée et le bruit typique de ces périphériques d'entrée donnent un retour immédiat que l'action a bien été envoyée au système.

Ce type de feedback permet de donner à l'utilisateur une indication sur la réponse temps-réel du système aux actions de l'utilisateur. Nous déclinons ce type feedback en deux catégories : les retours directs et les retours par reconnaissance. Il s'agit de distinguer les interactions qui nécessitent une interprétation de la part du système de celles qui ont un effet immédiat, les interactions discrètes des interactions continues. Tracer des lignes dans une application de dessin ou des caractères dans une application d'écriture ne nécessite pas d'interprétations algorithmiques immédiates de la part du système qui peut fournir alors un retour visuel direct en affichant en temps réel le tracé réalisé par l'utilisateur. Au contraire, manipuler un objet graphique avec plusieurs points de contact requiert une interprétation de la part du système afin de déterminer les intentions de l'utilisateur (mouvements parasites et mouvements intentionnels) en fonction des manipulations permises par l'objet. Dans les deux cas, le feedback visuel est déterminant puisqu'il indique à l'utilisateur qu'il a compris l'objectif des interactions. Dans le cadre du développement de l'intergiciel multi-tactile, nous avons tenu compte des contraintes de latences (cf. chapitre V - 3.2.) issues de la configuration du dispositif et des algorithmes de reconnaissance ou de lissage qui ont un impact considérable sur la qualité du feedback ressentie par les utilisateurs.

6.5. Gestion des erreurs

L'erreur n'est-elle pas humaine ? Lors de la communication entre deux personnes, il est rare qu'une conversation ne se déroule convenablement sans qu'un des deux interlocuteurs ne soit obligé de répéter, reformuler une phrase ou donner plus de précisions afin d'éviter toute erreur de compréhension. Norman particularise la communication homme-machine en ce sens où les interfaces n'ont pas la même tolérance. « *Appuyer sur le mauvais bouton peut entraîner le désordre.*¹⁹⁴ » Il distingue deux catégories d'erreurs : les étourderies (*slips*) et les fautes (*mistakes*). Les étourderies sont le résultat de comportements automatiques, lorsque l'inconscient détourne nos actions de leur intention initiale. Les fautes sont au contraire la conséquence d'actions conscientes et délibérées qui, guidées par un mauvais

¹⁹⁴ « Push the wrong button, and chaos may result. » In NORMAN Donald Arthur. *The Design of Everyday Things*. Basic Books, 2002, 272 p. Traduction libre Coquelin Marine et Morin Rudy.

jugement ou une interprétation incorrecte du fonctionnement d'un système, entraîne l'utilisateur à commettre des erreurs.

Les étourderies peuvent prendre différentes formes :

- *Les erreurs de capture* surgissent lorsqu'une activité, fréquemment réalisée prend soudainement la place d'une autre. Cela arrive lorsque deux actions ont en commun l'étape initiale.
- *Les erreurs de description* sont le résultat d'actions qui ont en commun d'autres actions similaires. Par manque de précision, l'action de l'utilisateur peut être correcte mais se porter sur le mauvais objet.
- *Les erreurs automatiques* sont le résultat de comportements instinctifs.
- *Les oublis*. C'est tout simplement lorsque l'utilisateur oublie tout ou une partie des actions à réaliser.
- *Les erreurs modales*. Ces erreurs peuvent survenir par manque de visibilité du mode actuel du système. L'utilisateur peut ne pas se souvenir des actions à réaliser dans le mode actif.

Il convient au designer de rendre visible les actions possibles et de mettre en place un système de feedback pour éviter, détecter, corriger ou annuler ces erreurs lorsqu'elles interviennent. Les fautes sont le résultat des mauvais choix de l'utilisateur. Une personne peut prendre la mauvaise décision, ne pas comprendre la situation qui lui est proposée ou ne pas prendre en compte tous les facteurs nécessaires. Les erreurs sont parfois liées à un problème de mémoire étant donné que la prise de décision est largement guidée par nos expériences passées. Norman avance quatre règles à suivre pour éviter ou corriger les erreurs utilisateur :

- Comprendre la cause des erreurs et concevoir des systèmes pour les minimiser.
- Rendre possible l'annulation d'actions ou mettre davantage en évidence celles qui ne peuvent être renversées.
- Faciliter la compréhension des erreurs qui interviennent et faciliter leur correction.

- Changer d'attitude. Penser que l'utilisateur fait l'exploration d'une interface par tâtonnements. Il s'agit de comprendre les actions des utilisateurs non pas en terme d'erreurs mais en terme d'approximations.

Ces considérations nous mèneront lors de la conception de l'intergiciel, à envisager un système d'aide didactique qui prendra en compte la notion d'apprentissage, d'actions répétées et de correction d'erreur. Ce que nous tenterons d'éviter également, c'est tout sentiment de culpabilité ou d'incompréhension de l'utilisateur à l'égard du système. Dans le design d'interaction des applications que nous développerons, nous mettrons également à contribution les méthodes de gestion des risques énoncés par Norman : les « inter-verrouillages » (*interlocks*) pour forcer les actions à se dérouler selon un ordre précis, les verrouillages internes (*lockin*), pour prévenir l'annulation prématurée d'une action lancée par l'utilisateur et les verrouillages externes (*lockout*), pour empêcher l'utilisateur d'atteindre des zones à risque ou empêcher certains événements de se produire.

CHAPITRE V PROTOTYPES

1. Prototypage

Les efforts que nous avons investis dans le développement et la construction de dispositifs interactifs multi-tactiles répondent aux besoins formulés par l'entreprise Virtual-IT, partenaire du CIFRE, qui a souhaité, dès mon arrivée, s'ouvrir à un nouveau domaine d'activité en créant un pôle de recherche et développement sur le développement de solutions interactives multi-tactiles. Dans un premier temps, nous avons réalisé une veille technologique sur les dispositifs et les environnements de développements existants, sur les techniques utilisées et les brevets déjà déposés dans le domaine. Ceci nous a permis de mettre en relief les contraintes techniques et financières qui nous ont conduits à préférer construire nos propres prototypes.

En effet, nous étions au début de l'année 2008 et les technologies tactiles multi-points, déjà commercialisées, souffraient pour la plupart de contraintes techniques majeures. Les cadres optiques infrarouges (*NextWindow*, *PQLabs*) revendiquaient un support multi-points mais ces technologies n'étaient capables d'en gérer correctement que deux simultanément. Hors, une de nos contraintes était de pouvoir développer et tester un environnement de développement multi-tactile offrant la liberté de créer des interactions manuelles et multi-utilisateurs plus complexes dont les modalités pouvaient impliquer jusqu'à une trentaine de points simultanés (trois utilisateur). Lorsque la table interactive *Microsoft Surface* a été commercialisée, fin 2008, nous étions intéressés par cette solution mais il nous est très vite apparu que la configuration de l'ordinateur intégré au dispositif était trop limitée pour faire tourner les applications en 3D temps réels que nous comptions développer. Enfin, le coût élevé de solutions plus ouvertes, comme l'écran de projection multi-tactile commercialisé par *Perceptive Pixel*, nous a définitivement convaincu de l'intérêt que nous avions de construire nos propres prototypes.

Au cours des deux premières années, nous avons travaillé sur la création d'un prototype fonctionnel et sur nos premières applications selon les directions définies par la cellule commerciale de l'entreprise. Nous avons ainsi dévoilé à nos clients les plus proches (DCNS, STEREAU), des applications multi-tactiles, en créant du contenu spécifique pour leur domaine d'application. Ceci nous a permis d'obtenir nos premiers retours utilisateurs. Ce n'est qu'à partir de la troisième année, grâce à l'expérience acquise en termes d'usabilité,

d'expression des besoins utilisateur et d'affinage technologique, que l'entreprise a décidé de commercialiser les produits tactiles issus du pôle R&D. Enfin, tout au long de ces trois années, nous avons participé à différents salons technologiques (voir Annexes - 3.) comme *Laval Virtual*, *Euronaal* et *Imagina*, en confrontant nos développements logiciels à un public plus large.

1.1. Prototype d'écran large multi-tactile

Bien que les tests de développement d'applications multi-tactiles fussent rendus possibles par l'utilisation d'outils de simulation de surface interactive compatibles avec le protocole TUIO (protocole d'objet tangible générique), il devint rapidement nécessaire de construire un dispositif multi-tactile afin de commencer le développement de l'intergiciel et des interfaces de programmation (*Application Programming Interface* ou API). Ce dispositif nous a permis de soumettre nos applications préliminaires aux utilisateurs pour évaluer en conditions réelles la pertinence de notre modèle d'interaction et pour valider les concepts de notre environnement de développement.

En prenant en considération les besoins exprimés par les clients de l'entreprise, le périmètre d'étude du projet de recherche et les prérequis techniques nécessaires au bon fonctionnement des applications, nous avons conçu un écran large multi-tactile répondant aux exigences suivantes :

- Une taille d'écran suffisamment grande pour permettre des interactions multi-utilisateurs (2 x 1,50m).
- Une qualité d'image satisfaisante en termes de résolution (1920 x 1080px), de luminosité et de contraste, afin de supporter une vision rapprochée.
- Un écran mat pour éviter les reflets.
- Une surface d'écran lisse et non adhérente pour augmenter la qualité des contacts.
- Une technologie tactile permettant des interactions riches et multi-utilisateurs (nombre de points de contact, précision).
- Un dispositif facilement démontable et transportable pour la participation à des salons professionnels.

- Une configuration optique et électronique peu sensible aux interférences lumineuses (lumière ambiante, changements de luminosité).
- Le choix de matériaux et d'une technique d'illumination infrarouge répondant aux normes de sécurité pour la mise à disposition au public.
- Un faible coût de production.



Figure 62 - Vue avant du prototype T1 multi-tactile. L'écran a une largeur de 2m et une hauteur de 1,5m. Aujourd'hui, ce prototype est installé à l'école Epitech Toulouse dans le cadre du partenariat entre l'école et l'entreprise.

1.2. Technique d'illumination

Le choix d'une technique d'illumination est déterminant puisqu'il influence considérablement le type d'écran, de caméra et de matériaux utilisés. Nous présentons au chapitre III - 2.2. les différentes solutions qui ont retenu notre attention. Afin d'évaluer ces différentes techniques d'illumination et pour désigner celle qui correspondait le plus à nos besoins, nous avons fabriqué des dispositifs miniatures, d'une surface de 29,7 x 42cm, pour valider leurs défauts et leurs qualités. Voici les résultats de cette étude :

L'illumination par réflexion totale frustrée - FTIR (principe décrit Chapitre III - 2.2.)

Popularisée par les travaux de recherche de Jeff Han¹⁹⁵, cette technique nous a séduit car elle était peu encombrante et tenait la promesse, contrairement aux autres techniques d'illumination, de permettre une illumination du doigt uniquement lorsque celui-ci entre réellement en contact avec la surface. Les tests que nous avons réalisés, avec le concours d'un ingénieur en électronique, nous ont permis d'accéder à de très bons résultats, en termes de luminosité et de contraste, en complétant les caractéristiques techniques définies par Jeff Han :

- La plaque d'acrylique ne peut être inférieure à 6mm et doit avoir les tranches polies,
- Il est indispensable de masquer les bords de l'écran, là où sont placés les diodes électroluminescentes, afin d'éviter toute fuite de lumière,
- La longueur d'onde de lumière, émise par les diodes, donne de meilleurs résultats lorsqu'elle est située entre 780 et 850nm,
- Pour la diode, une intensité radiante supérieure à 120 mW est préférable,
- Utiliser les diodes en lumière pulsée permet d'augmenter la luminosité et la portée de l'illumination,
- L'angle d'émission des diodes doit être situé entre 96° et 120° pour assurer le bon fonctionnement du principe de réflexion totale frustrée,
- Plus la surface à couvrir est grande et plus les diodes doivent être proches les unes des autres.

Malgré les bons résultats obtenus avec ces réglages, nous avons été confrontés à un problème assez ennuyeux puisque nous nous sommes rendu compte que le principe de réflexion totale frustrée fonctionnait considérablement moins bien lorsque l'utilisateur avait les doigts très secs. Une solution existait à ce problème, consistant à recouvrir la surface en polycarbonate d'une couche intermédiaire, une surface de médiation (*compliant surface*), en silicone ou toute autre matière permettant de provoquer le principe de réflexion totale frustrée. Toutefois, la taille considérable de l'écran que nous devons couvrir et les difficultés

¹⁹⁵ HAN Jefferson Y. *Low-cost multi-touch sensing through frustrated total internal reflection*. In *Proceedings of the 18th annual ACM symposium on User interface software and technology (UIST '05)*. ACM, New York, USA, 2005, pp. 115-118

techniques de l'application de la surface de médiation, nous ont découragé de poursuivre dans cette direction.

Plan de lumière laser - LLP (principe décrit Chapitre III - 2.2.)

Cette technique d'illumination nous a également séduits car elle est peu encombrante et elle permet de couvrir de grandes surfaces. De plus, la bande passante du spectre lumineux étant très faible puisque l'illumination se fait grâce à des lasers, cela augmente l'efficacité de la détection des points de contact tout en diminuant l'influence des interférences lumineuses. Toutefois, deux facteurs nous ont contraints de mettre de côté cette solution.

D'une part, afin de créer un plan de lumière homogène sur l'ensemble de la surface à couvrir, les tests que nous avons menés ont montré qu'il fallait installer au minimum huit modules lasers d'une puissance de 20mW chacun. Installer ces lasers sur des fixations en bordure d'écran requiert une extrême précision puisqu'une légère modification de l'inclinaison d'un laser, d'un côté de l'écran, peut entraîner un décalage de plusieurs centimètres du plan de lumière à l'extrême opposé. Reproduits sur les huit lasers, ces décalages entraînent une répartition hétérogène de la lumière en plans non parallèles et cela diminue la qualité de la détection des points de contact, qui peuvent alors être détectés bien avant le contact réel du doigt sur l'écran.

D'autre part, même si les lasers que nous comptions utiliser étaient munis d'une lentille génératrice de ligne, réduisant ainsi son intensité, nous ne souhaitons pas courir le risque lié à l'utilisation de lasers infrarouges. Lors de nos tests, nous portons d'ailleurs des lunettes de protection spécialisées car la lumière laser peut causer des dommages irréversibles aux yeux. De plus, étant invisible, la lumière infrarouge ne provoque pas le phénomène d'éblouissement qui peut dans certains cas obliger le regard à se détourner de la source laser. Comme nous souhaitons présenter ce dispositif au public, même si les risques étaient limités par la pose de caches optiques, un accident était toujours possible et nous n'avons pas souhaité en courir le risque.

Illumination diffuse - DI (principe décrit Chapitre III - 2.2.)

Il s'agit de la technique d'illumination que nous avons sélectionnée car elle est plus simple à installer et à configurer que les techniques évoquées précédemment. Son fonctionnement, par diffusion de la lumière infrarouge à travers une surface diffusante, n'impose pas le

recours à une surface de médiation. L'avantage de cette configuration réside en l'utilisation d'une surface unique, qui sert à la fois d'écran et de diffuseur pour la lumière infrarouge. De plus les principes optiques impliqués dans cette configuration autorisent la pose d'un verre de protection devant la surface afin de la protéger du vandalisme et renforcer la rigidité de l'écran. Toutefois, cette technique d'illumination génère des blobs (zone de lumière émise par le doigt en contact) moins contrastés et implique l'installation encombrante d'illuminateurs infrarouges derrière l'écran de projection.

Nous avons donc travaillé à réduire ces défauts grâce au savoir-faire de Jean-François Montchamp, ingénieur en optique et en électronique, qui a participé au projet. D'une part, nous avons réalisé une batterie de tests sur une dizaine de surfaces en polycarbonate sablées afin de sélectionner celle qui répondrait aux exigences suivantes :

- Un taux de diffusion élevé afin de limiter la zone de détection des doigts à l'avant de la surface.
- Un taux de réflexion minimal, à l'arrière, afin de limiter la pollution lumineuse générée par la réflexion éventuelle des illuminateurs.
- Une diffusion homogène de la lumière, pour des angles d'incidence compris entre 70° et 75°.
- Un taux de luminosité et de contraste élevé pour la vidéo-projection.
- Une surface suffisamment sombre et opaque pour limiter les phénomènes de points chauds (*hotspots*), zone de l'écran plus lumineuse créée par la vue de la source de lumière du vidéoprojecteur à travers l'écran.
- Une épaisseur suffisamment importante pour ne pas subir de déformations lors du contact.

D'autre part, comme les illuminateurs infrarouges vendus dans le commerce ne correspondaient pas aux caractéristiques que nous avons formulées, nous avons décidé de construire nos propres illuminateurs. Les tests d'illumination préliminaires nous ont aidés à définir les spécifications nécessaires pour obtenir des zones de contact lumineuses et contrastées :

- Une longueur d'onde située entre 790 et 850nm correspondant à la bande passante autorisée par la caméra.

- Un angle d'émission suffisamment large pour diminuer la distance entre les illuminateurs et les écrans et obtenir une répartition plus homogène de la lumière infrarouge.
- Une intensité variable pour ajuster l'illumination globale.
- Un système électronique de lumière pulsée pour profiter de la puissance des diodes CMS.
- Un système de sécurité (coupure automatique) et de refroidissement (ventilateur au dos) si la température excède les conditions normales d'utilisation.

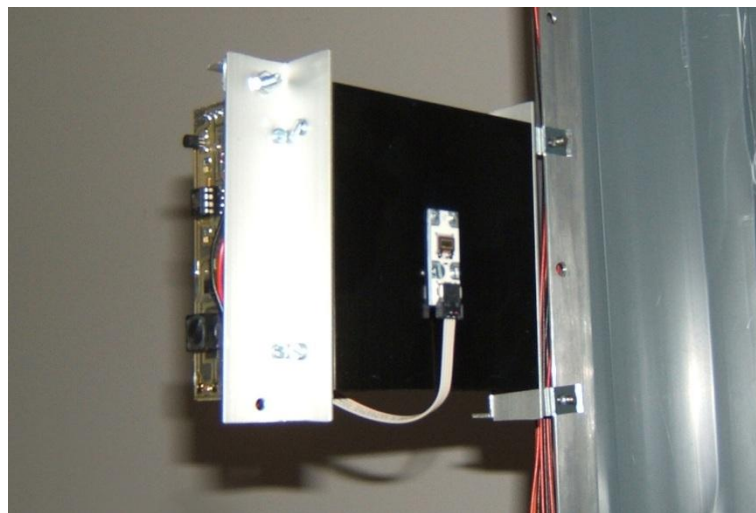


Figure 63 - Illuminateur infrarouge conçu en interne. 16 illuminateurs de ce type, situés à une cinquantaine de centimètres tout autour de l'écran, nous ont permis d'obtenir une illumination homogène et d'intensité suffisante. Les schémas électroniques sont soumis à la politique de confidentialité de l'entreprise.

1.3. Système de capture optique

La configuration technique que nous avons choisie implique l'utilisation d'une caméra dont le rôle est de fournir au système des images, en temps-réel, qui sont ensuite analysées par différents algorithmes de détection et de suivi de points. Le système optique que nous avons élaboré est constitué d'une caméra sensible aux infrarouges, d'un filtre optique passe-bande et d'une lentille à focale variable dépourvue de filtre (anti)infrarouge (cf. figure 64).



Figure 64 - Configuration du système de capture optique.

Les études théoriques et les mesures effectuées sur les bancs d'essai, nous ont permis de mettre en évidence des facteurs déterminants pour assurer la qualité de la détection des points de contact :

- *La résolution du capteur* de la caméra a une influence considérable sur la détection des doigts puisqu'un des filtres du traitement d'image analyse la « netteté » d'un blob pour déterminer la distance du doigt par rapport à l'écran. D'autre part, dans le cas de la reconnaissance de marqueurs, une résolution plus élevée facilitera sa reconnaissance. Dans notre configuration, nous avons estimé qu'une résolution de 640 x 480 pixels était suffisante. Nous avons envisagé d'utiliser une résolution plus élevée ou de développer un système de captation multi-caméras mais les coûts de calcul processeur pour traiter de telles images et le temps de développement des algorithmes de jointure (*stitching*) nous ont contraint de garder cette configuration pour le prototype.
- *La cadence d'image* de la caméra correspond au nombre d'images par seconde que la caméra est capable de capturer. Plus la cadence d'image est élevée, plus le système est réactif aux actions de l'utilisateur. L'analyse du mouvement d'un point de contact par le système se produit en comparant les coordonnées d'un point par rapport aux coordonnées des points de l'image précédente. En réduisant l'écart temporel entre deux images, on réduit la latence des réponses systèmes (feedback), aux mouvements utilisateur. Nous avons estimé qu'une cadence de 60 images par

seconde était un bon compromis entre la qualité de la détection et les capacités des ordinateurs que nous utilisons.

- *L'interface de la caméra* correspond à la norme ou le bus utilisé pour la connexion entre la caméra et le système. Les caméras sont généralement compatibles avec l'USB (*Universal Serial Bus*) et le FireWire (IEEE 1394). Ce dernier est préférable puisqu'il bénéficie d'une latence moins importante et contribue à augmenter la réactivité générale du système.
- *La lentille optique* dépend essentiellement de deux facteurs : la taille de la surface à capturer et la distance de la caméra par rapport à cette surface. En fonction de ce rapport et de la taille du capteur de la caméra, la longueur focale de la lentille peut être déterminée. Plus la distance focale est courte, plus l'angle de vue est important mais peut entraîner une distorsion optique en barillet. Ce phénomène optique peut toutefois être corrigé de manière logicielle.
- *Le filtre optique passe-bande*, ajouté à la caméra, permet de filtrer une partie du spectre électromagnétique. D'une part les longueurs d'onde de la lumière visible et d'autre part, les longueurs d'onde qui ne correspondent pas au spectre d'émission des illuminateurs infrarouges. Il s'agit d'un premier traitement optique qui permet à la caméra de capturer l'image des objets éclairés par les illuminateurs, qui apparaissent de manière contrastée par rapport aux objets éclairés par la lumière ambiante (éclairage ambiant et solaire). En fonction de la réponse spectrale du capteur et de l'émission spectrale des diodes, nous avons déterminé la fréquence du filtre passe-bande, fixée à 850nm. Sa bande passante est considérablement déterminée par les conditions de lumière et la pollution lumineuse de l'environnement. C'est pourquoi nous avons à notre disposition des filtres de largeurs spectrales différentes : 10nm, 30nm et 40nm.

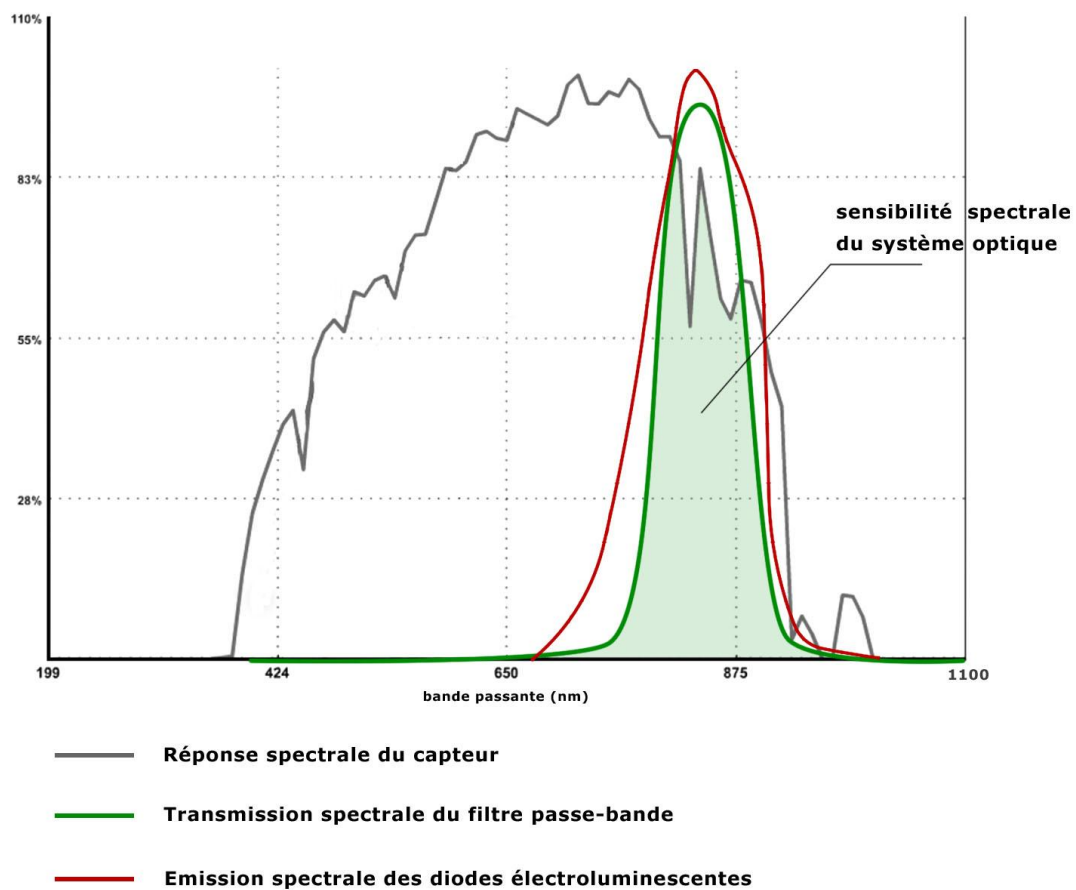


Figure 65 - Sensibilité spectrale du système optique utilisé pour le prototype.

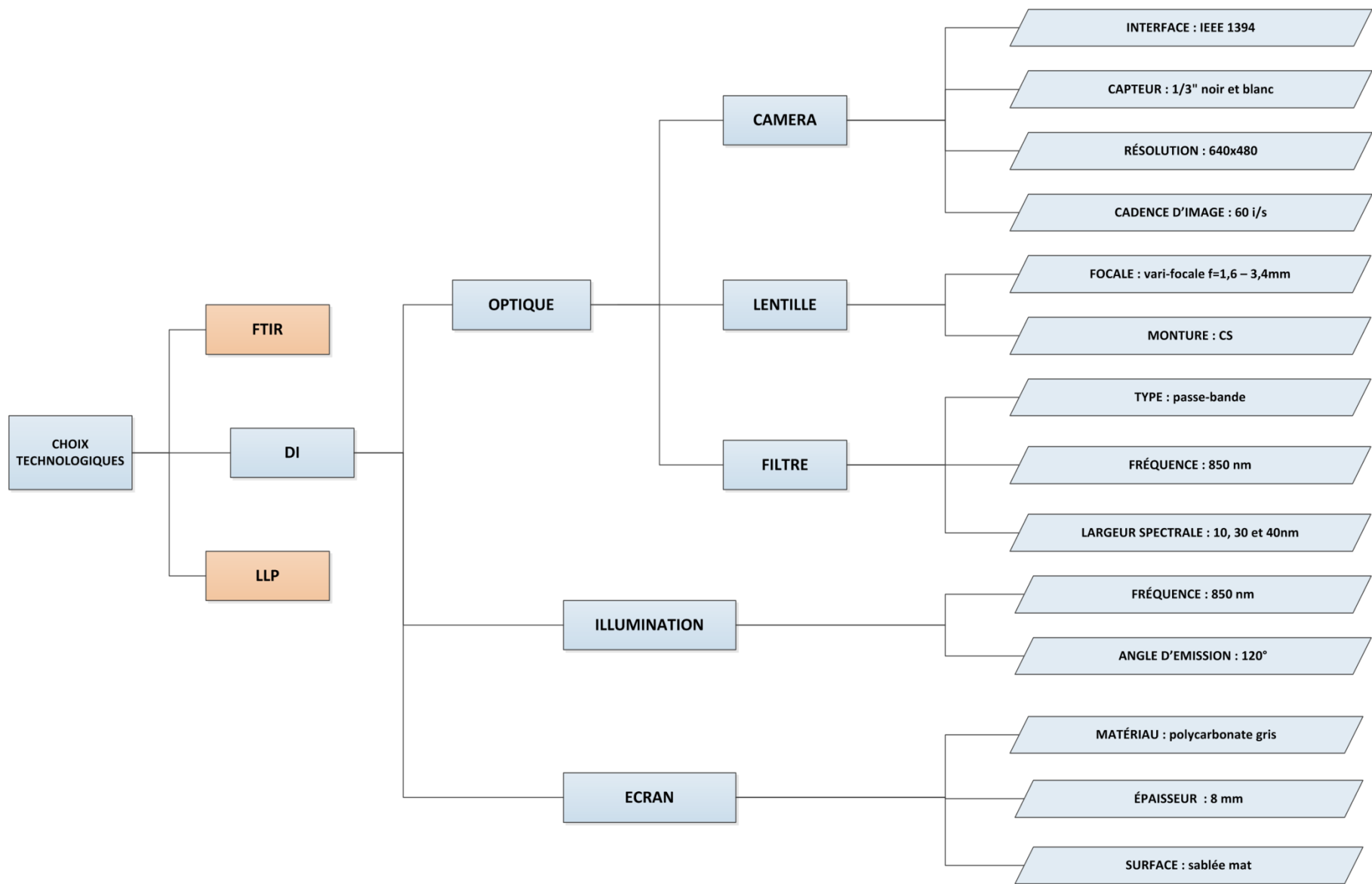


Figure 66 - Arbre des choix technologiques

2. Prérequis logiciels

2.1. Vision par ordinateur

La configuration d'illumination diffuse que nous avons sélectionnée nous a permis d'obtenir des images contrastées des zones de contact générées par la pose des doigts sur l'écran. Ces zones plus lumineuses, appelées *blobs*, sont capturées par la caméra qui envoie au système des images monochromes, traitées par différents algorithmes de filtrage, de détection et de suivi (*tracking*) de points. Pour alléger la charge de travail et concentrer nos efforts sur le développement d'un environnement de développement (*framework*) multi-tactile, nous avons décidé d'utiliser un logiciel de traitement d'image appelé *CCV* (*Community Core Vision*). *CCV* est une solution logicielle libre et multiplateforme de vision par ordinateur. Elle prend en entrée le flux vidéo envoyé par la caméra et envoie en sortie les données (position, taille des blobs, etc.) et les événements (doigt posé, déplacé ou enlevé) issus du suivi de déplacement des points de contact. Les données analysées sont ensuite encodées dans un protocole appelé *TUIO*, dont le format de données contient toutes les informations nécessaires, sous la forme de données brutes (*raw data*), pour être converties par l'intergiciel que nous avons développé. *CCV* est l'évolution du logiciel *tbeta*, projet développé et maintenu par la communauté *NUI Group* pour favoriser le développement de projets multi-tactiles. Ce logiciel utilise en partie l'environnement de développement *OpenFrameworks* et la librairie de traitement d'image *openCV* (*Open Source Computer Vision*) sous licence BSD.

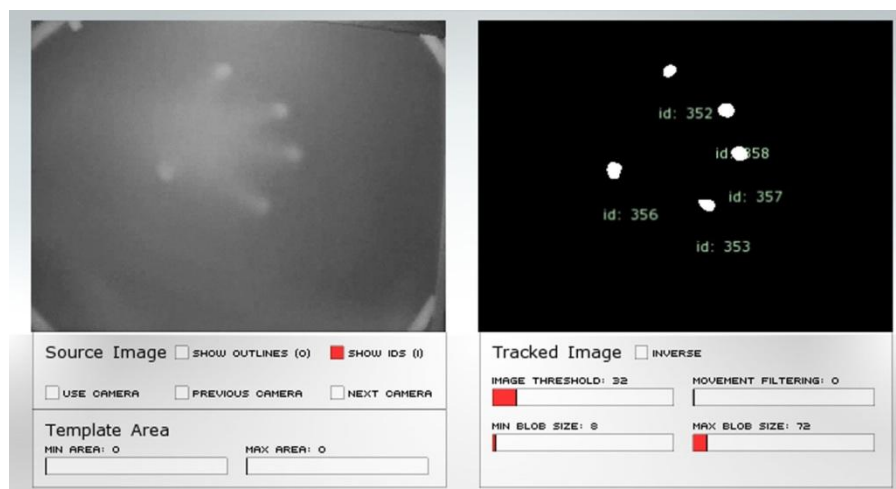


Figure 67 - Captures d'images du logiciel CCV. A gauche, l'image brute issue de la caméra. A droite, l'image après traitement et suivi de points.

Avant d'appliquer l'algorithme de suivi de points, l'image passe à travers une succession de filtres qui permettent de « nettoyer » l'image pour ne laisser visibles que les zones de contacts. L'image issue de la chaîne de filtres est monochrome (pixels blancs correspondent aux zones de contact) et permet à l'algorithme de suivi de point de traiter l'image avec une meilleure précision. Les filtres disponibles sont :

- Suppression de fond (*background subtraction*): effectue une comparaison entre l'image de référence et l'image actuelle afin de supprimer toute information de pixel identique d'une image à l'autre.
- Filtre de lissage : lisse l'image et supprime le bruit vidéo.
- Filtre passe-haut de netteté : supprime les parties floues de l'image et conserve les parties les plus nettes.
- Filtre d'amplification : augmente la luminosité des pixels peu lumineux.

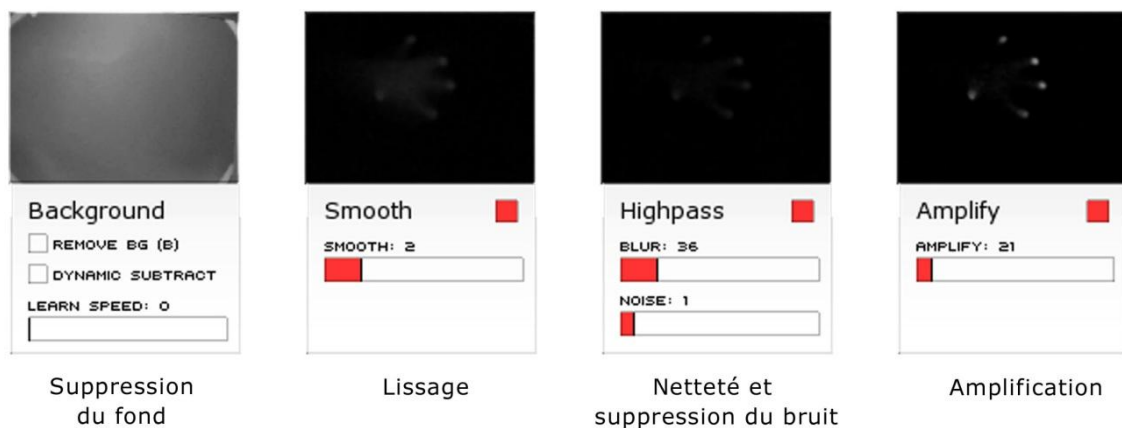


Figure 68 - Capture d'image du logiciel CCV. Chaîne de filtre appliquée à l'image brute issue de la caméra.

2.2. Détection et suivi des points

Pour chaque nouvelle image envoyée par la caméra au système, le programme utilise la fonction *cvFindContours()*, de la librairie *OpenCV*, pour obtenir une liste des contours présents dans l'image. Chaque contour, ainsi que ses propriétés, est stocké dans un tableau puis analysé pour déterminer si la zone de contact correspond à un marqueur posé sur l'écran ou à un doigt. Cette distinction est réalisée par comparaison entre le contour détecté et un polygone à quatre côtés et par le calcul des angles de chaque coin. Si le contour

possède quatre côtés et si les angles font 90°, l'application considère qu'il s'agit d'un marqueur. Si le polygone est plus complexe, elle considère qu'il s'agit d'un doigt. En juxtaposant une ellipse sur le contour, l'application détermine la position, l'orientation et la taille de la zone de contact. Si la taille des blobs appartient à l'encadrement de valeurs défini par les seuils réglés dans l'application, les zones détectées sont ajoutées à la liste des points de contact tactiles.

Suite à l'identification de toutes les zones de contact, le logiciel CCV applique un algorithme de suivi de points. Afin de limiter la charge processeur nécessaire au suivi de points de contact simultanés, dont la quantité peut parfois excéder une trentaine de points, les développeurs ont préféré écarter les approches traditionnelles comme le filtre kalman. Etant donné que le suivi de points s'applique sur des images peu espacées dans le temps (33 ms ou 16ms généralement), et sachant qu'un doigt ne peut parcourir qu'une courte distance en si peu de temps, le modèle de suivi utilisé est l'approche des k plus proches voisins. L'algorithme des plus proches voisins consiste à passer sur l'ensemble des points d'un ensemble et à vérifier si ce point est plus proche ou non qu'un des plus proches voisins déjà sélectionné, et si oui, l'insérer.

Afin de déterminer les relations entre les coordonnées des points issus de l'image précédente et ceux de l'image actuelle, CCV crée une matrice qui contient tous les états de conversions possibles. L'état d'un point correspond soit à un des points déjà existants, soit à un nouveau point. Avec n représentant le nombre de blobs actifs dans l'image précédente et a le nombre de blobs actifs dans l'image courante, le nombre d'états possibles E est calculé selon l'équation suivante :

$$E = \frac{(n + m)!}{(n + m - a)! m!}$$

Où le nombre de nouveaux points est $m = a - n$

Exemple :

Soit p1, p2 et p3, trois points de contact existants dans l'image précédente et p4 un nouveau point de contact dans l'image courante. En considérant que p1, p2 et p3 ont pu changer de position entre les deux images, le nombre possible d'états E peut être calculé de la manière suivante :

$$E = \frac{(3 + 1)!}{(3 + 1 - 4)! 1!} = 4! = 24$$

CCV recherche ensuite l'ensemble d'états qui contient la plus petite valeur. Cette valeur est déterminée en ajoutant les distances de l'ensemble des probabilités de chaque état, entre les points de l'image précédente et ceux de l'image actuelle selon la formule de géométrie analytique suivante :

$$d = \sqrt{(x_{p'} - x_p)^2 + (y_{p'} - y_p)^2}$$

La formule est appliquée à l'ensemble des états possibles pour trouver celui où les points sont les plus proches. Ainsi il est possible de faire un suivi des points de contact et d'identifier les nouveaux. Dans le cas où deux distances partagent la même valeur, CCV choisira par défaut la première. Cette situation peut surgir lorsque deux points sont très proches et peut avoir comme conséquence l'échange d'identifiant des deux points concernés. Dans ce genre de situation, nous regretterons que CCV n'utilise pas de filtre prédictif comme le filtre Kalman qui aurait permis d'identifier un point en fonction de sa trajectoire et de sa vitesse.

2.3. Protocole TUIO

Tuio, pour objet d'interface utilisateur tangible (*Tangible User Interface Objects*), est le protocole par défaut utilisé par le logiciel de vision par ordinateur CCV que nous avons utilisé. Depuis la publication des spécifications du protocole en 2005 par *Martin Kaltenbrunner et al.*¹⁹⁶, dans le cadre du projet *reacTable*, il a été adopté dans de nombreux

¹⁹⁶ KALTENBRUNNER Martin, BOVERMANN T., BENCINA R., Costanza E. *TUIO - A Protocol for Table Based Tangible User Interfaces*. In *Proceedings of the 6th International Workshop on Gesture in Human-Computer Interaction and Simulation*, Vannes, France, 2005

développements d'interfaces tangibles et multi-tactiles, dans un contexte universitaire ou commercial. Ce protocole définit une interface de communication générale et flexible entre les périphériques d'entrée utilisateur multi-tactiles et les applications. Lorsqu'un contact est effectué par un objet, un profil de type « *object* » contient l'identifiant du marqueur détecté, sa position et son orientation sur la surface. Dans le cas d'un contact effectué par le doigt, un profil de type « *cursor* » est créé, contenant un identifiant, une session d'identification unique et les informations relatives à sa position et à ses états (événements).

Ce protocole est encodé au format OSC (*Open Sound Control*), protocole ouvert de transport de paquets utilisé pour la communication entre ordinateurs, synthétiseurs et autres appareils multimédia. Ce format permet l'encodage binaire de données sous la forme d'un message normalisé. Le protocole TUIO implique le transport des messages OSC en les encapsulant dans des paquets UDP (*User Datagram Protocol*) afin de gagner en simplicité et en vitesse. La version 1.1 du protocole offre de nouvelles méthodes de transport, par TCP (*Transmission Control Protocol*) ou FLC (*flash local connection*) mais le protocole n'est pas encore à jour dans tous les programmes que nous utilisons.

Le protocole TUIO définit deux classes de messages :

- Les messages *SET* sont utilisés pour communiquer des informations sur l'état d'un objet comme sa position et son orientation.
- Les messages *ALIVE* indiquent la liste actuelle des objets présents sur la surface en utilisant une liste d'identifiants de session uniques.
- Les messages *FSEQ* permettent d'identifier chaque mise à jour de valeur par un identifiant unique d'image.

2.4. TuioSimulator

Tout au long du développement de l'intergiciel et des applications multi-tactiles, nous avons utilisé un simulateur d'interface tangible pour faciliter les tests et accélérer les temps de débogage. Nous avons choisi *TuioSimulator*, logiciel en java développé par l'équipe à l'origine du logiciel open source *reactiVision*, pour sa stabilité et les multiples fonctionnalités proposées par le logiciel. *TuioSimulator* simule une table interactive virtuelle qui détecte et analyse les propriétés des objets ou des points tactiles générés par les clics de la souris. Il

implémente le protocole TUIO et permet de spécifier l'adresse et le port du serveur de l'application tactile. Il est possible, grâce à ce simulateur, de générer plusieurs points de contact et de modifier leur position individuellement ou de manière simultanée. Enfin, il est possible de l'utiliser à travers le réseau local en spécifiant l'adresse de la machine distante.

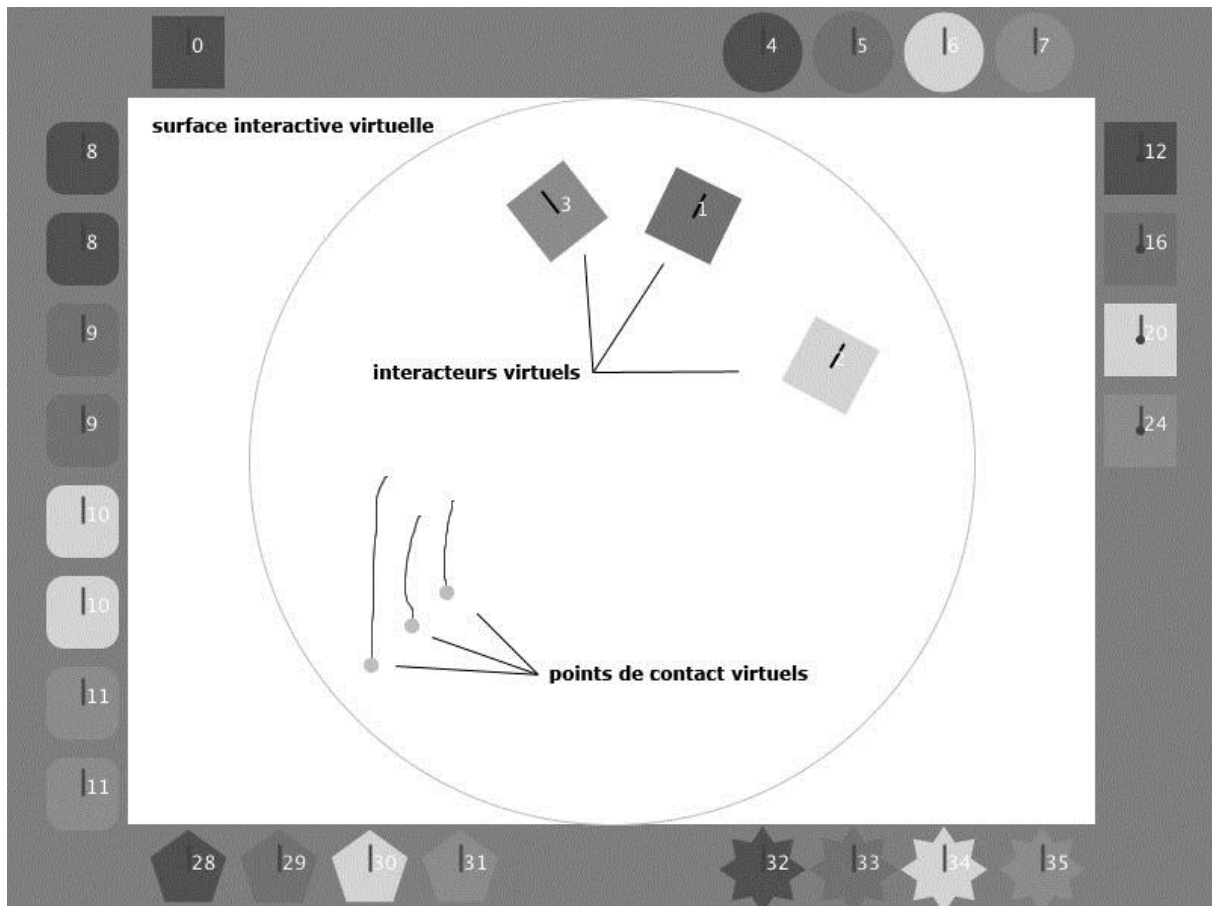


Figure 69 - Capture d'écran de l'interface de *TuioSimulator*.

3. Considérations techniques

3.1. Déformation optique

Pour la construction du prototype d'écran multi-tactile, la configuration que nous avons choisie utilise la caméra comme capteur. La précision de la détection des points de contact dépend de multiples facteurs comme la résolution et la sensibilité spectrale du capteur à la longueur d'onde utilisée par les illuminateurs. Elle dépend également de la calibration initiale (précision et nombre de points de référence) réalisée via l'interface du logiciel CCV.

Lorsque nous avons réglé la lentille optique de la caméra en fonction de la distance et de la taille de la surface, l'image issue de la caméra était déformée par l'effet de distorsion en barillet (connu sous le nom d'effet *fisheye*) causé par la distance focale relativement courte de la lentille.

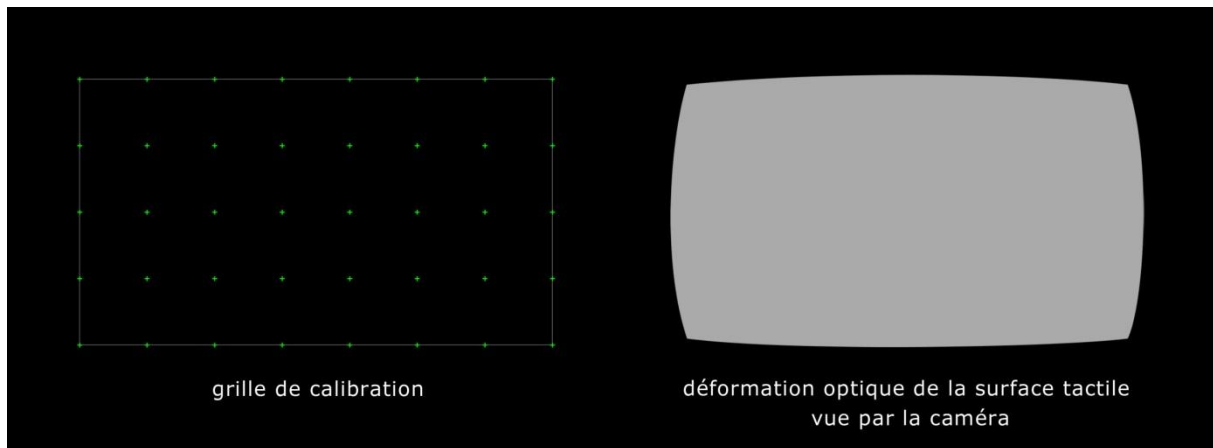


Figure 70 - Illustrations de la vue caméra. A gauche, la grille de calibration. A droite, la représentation de la zone écran déformée par la distorsion en barillet.

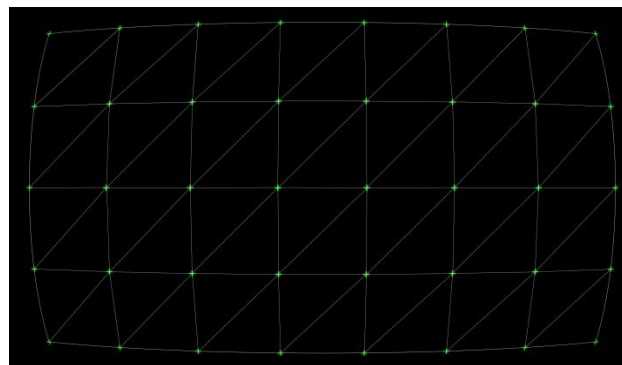


Figure 71 - Grille de calibration projetée sur la surface déformée.

Si nous faisons l'enregistrement d'un nombre insuffisant de points, la calibration ne permettrait pas au logiciel de faire une projection correcte étant donné la forte distorsion de l'image source. Toutefois, CCV dispose d'un système de correction de points basique qui s'est révélé suffisamment efficace pour les besoins du prototype. Pour cela, nous réalisons la calibration en enregistrant au minimum quarante points de contact. Pour corriger la déformation radiale causée par la lentille optique, CCV utilise la projection par coordonnées barycentriques.

Pour une grille rectangulaire donnée, en connaissant les coordonnées en deux dimensions (x, y) de chaque nœud de la grille, il est possible de déterminer les coordonnées projetées d'un point P en connaissant les coordonnées exactes de ses proches voisins référencés (cf. figure 72).

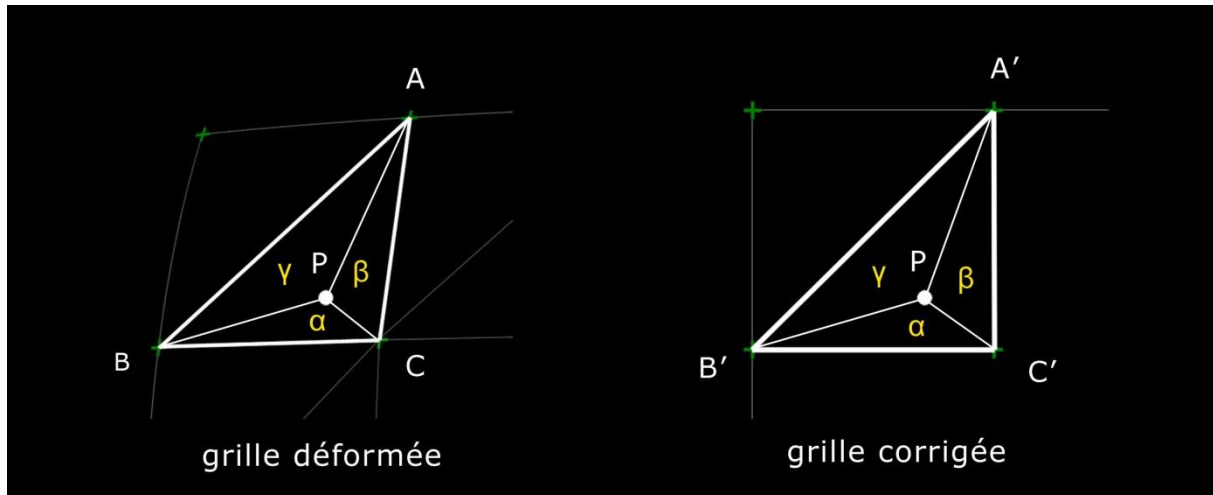


Figure 72 - Illustration de la correction barycentrique. A gauche, le triangle de la grille déformée. A droite le triangle de la grille corrigée. P correspond au point de contact situé en haut à gauche de l'écran.

En considérant qu'il existe suffisamment de triangles pour justifier une approximation de la distorsion non linéaire de la grille par une transformation affine, nous pouvons assumer que les coordonnées barycentriques α , β et γ du point P sont les mêmes dans la grille déformée que dans la grille de référence. Le point P peut être défini comme la somme des poids calculés selon les formules suivantes :

$$\alpha = \frac{\Delta PAB}{\Delta ABC} \quad \beta = \frac{\Delta PAC}{\Delta ABC} \quad \gamma = \frac{\Delta PBC}{\Delta ABC} \quad \alpha + \beta + \gamma = 1$$

A partir des coordonnées barycentriques, CCV calcule les coordonnées corrigées en utilisant les poids α , β et γ .

3.2. Latence du système

Dans le modèle théorique d'interaction multi-tactile que nous avons défini, nous avons insisté sur l'influence du feedback visuel des interactions utilisateur sur la « visibilité » de l'interface. La qualité du feedback visuel étant en grande partie liée à l'immédiateté du retour système, nous nous sommes inquiétés des effets de la latence causée par les périphériques d'entrée multi-tactiles. Lors de la conception du prototype, nous avons fait l'estimation de l'influence de la combinaison matérielle (caméra) et logicielle (traitement d'image) sur les performances du système pour réduire la latence générale éprouvée lors des premiers tests. En partant des travaux de thèse réalisés par Laurence Muller¹⁹⁷, nous avons mis en évidence le rôle joué par les différents éléments impliqués dans les processus de captation, de traitement et de diffusion du système.

Latence matérielle

La première étape du traitement d'image consiste à capturer l'image de la caméra. Selon les spécifications données par l'IIDC¹⁹⁸ (*Instrumentation and Industrial Digital Camera*), il est possible de calculer le temps de transfert d'une image par l'intermédiaire du bus IEEE1394 (*FireWire*). Le *FireWire* utilise en mode isochrone un multiplexage temporel où les paquets sont envoyés toutes les 125 microsecondes. Le format d'image monochrome (Format 0, Mode_5, 8 bit) dont nous avons besoin pour obtenir une qualité élevée de la détection des points de contact utilisait une résolution de 640 x 480 pixels à 60 images par seconde. Ce format suppose un transfert de 640 quadlets par paquets. 1 quadlet = 32 bits et peut donc coder 4 pixels. Chaque pixel de ce format étant codé sur 8 bits, donc le temps de transfert T théorique d'une image jusqu'au système est donc :

$$T = \frac{640 \times 480}{4 \times 640} \times 0.125 = 15 \text{ ms}$$

Les vidéoprojecteurs génèrent également un temps de latence, qui ne figurait pas dans les spécifications du produit¹⁹⁹ que nous avons utilisé. Toutefois, il nous a été possible de la

¹⁹⁷ MULLER Laurence Y. L. *Multi-touch displays : design, applications and performance evaluation*. Thèse, sous la direction de BELLEMAN R.G. Universiteit van Amsterdam, Amsterdam, 2008

¹⁹⁸ 1394 Trade Association. DC 1394-based Digital Camera Specification Ver.1.31. 2004. Disponible ici : http://damien.douxchamps.net/ieee1394/libdc1394/iidc/IIDC_1.31.pdf

¹⁹⁹ Marque du vidéoprojecteur sous la politique de confidentialité de l'entreprise.

mesurer grâce à l'outil développé par Laurence Muller²⁰⁰ appelé *Beamer Latency Measurement Tool*. Cet outil de mesure permet de comparer le délai en nombre d'images entre l'affichage du vidéoprojecteur et celui d'un écran traditionnel. Après avoir activé la synchronisation verticale de la carte graphique et lancé l'application, des barres verticales égales au nombre d'images par seconde supportées par les périphériques sont affichées sur l'écran traditionnel comme sur la surface de projection et une barre de référence avance d'une barre à l'autre à raison d'un cycle par seconde. En réalisant une photo des deux écrans, avec un temps d'exposition très court, il nous a été possible de comparer l'écart entre les deux périphériques. Pour notre test, les spécifications de l'écran LCD que nous avons utilisé indiquait une latence insignifiante de l'ordre de 2ms. En comparant la position de la barre de référence d'un écran à l'autre sur la photo, nous avons mesuré une différence d'une image. Le rafraîchissement de l'écran étant réglé sur 60 images par seconde, le délai d'affichage entre les deux écrans peut être calculé de la façon suivante :

$$\text{Latence} = \text{Cadence d'image} \times \text{différence d'images relevée}$$

$$\text{Latence} = \frac{1}{60} \times 1 = 0.0167 \text{ secondes} = 16,67 \text{ millisecondes}$$

En ajoutant le délai mesuré à la latence théorique de l'écran LCD, nous pouvons estimer la latence du vidéoprojecteur à 18,67 millisecondes (16,67 + 2).



Figure 73 - Représentation des mesures de latence effectuée sur le vidéoprojecteur et l'écran.

²⁰⁰ MULTIGESTURE.NET. Disponible sur : <http://www.multigesture.net>

Latence logicielle

Entre l'acquisition et la projection, les filtres de traitement d'image et les algorithmes de suivi de points agissent également sur la latence du système. Ces calculs étant exécutés par le processeur de l'ordinateur, nous avons privilégié l'utilisation d'une machine avec un processeur puissant. Pour déterminer la latence générée par les traitements d'image, nous nous sommes référés à la valeur indiquée par CCV. Nous avons relevé, pour l'ensemble des calculs nécessaires pour une image de 640 x 480 pixels au format 0, une utilisation CPU de 12% et un temps de latence de 17 millisecondes.

Latence générale

Les mesures de temps de latence ont été réalisées dans un contexte optimal, sans autre programme lancé simultanément sur l'ordinateur. Nous avons bien conscience du caractère théorique et approximatif des valeurs que nous avons obtenues mais elles rendent compte d'une latence qui est tout de même perceptible pour l'utilisateur.

En ajoutant les différentes valeurs obtenues, il est possible d'établir le temps de latence générale théorique de la configuration de notre prototype à 48,67 millisecondes.

$$15 + 16,67 + 17 = 48,67 \text{ ms}$$

Avec un temps de latence légèrement inférieur à $\frac{1}{20}$ seconde les conséquences sur le feedback visuel ne sont perceptibles que sur des mouvements rapides de la part de l'utilisateur. Lors du développement de l'environnement multi-tactile, nous veillerons à minimiser le temps de traitement des informations tactiles, en optimisant les algorithmes et les protocoles de communication, afin de ne pas augmenter la latence générale du système déjà conditionnée par les périphériques d'entrée et de sortie.

CHAPITRE VI ENVIRONNEMENT DE DÉVELOPPEMENT

« La meilleure façon de prédire l'avenir est de l'inventer. »

« The best way to predict the future is to invent it. »

Alan Kay

En juin 2008, nous disposions d'une première version du prototype multi-tactile, fonctionnel, qui nous a servi de support pour commencer le développement d'applications. Nous avons profité de ces premières réalisations pour faire l'inventaire des environnements de développements tactiles existants et procéder à leur évaluation. D'une part, en les mettant à l'épreuve du modèle d'interaction que nous avons défini préalablement et d'autre part, en confrontant la viabilité de ces solutions aux perspectives de développement de l'entreprise sur le long terme.

Le programme de recherche et développement de l'année 2009 a été partagé entre l'amélioration du prototype (précision, sensibilité, encombrement), l'évaluation de nouveaux périphériques (cadres optiques *NextWindow* et *PQLabs*) et le développement, interne comme externe d'applications multi-tactiles. Cette phase a été déterminante car elle nous a permis d'obtenir les premiers retours utilisateur et de percevoir les limitations des architectures logicielles existantes. Rapidement, nous avons compris la nécessité de créer notre propre environnement de développement si jamais nous souhaitions proposer des solutions multi-tactiles aux clients de l'entreprise. Ce n'est qu'au bout d'une année d'observation, d'analyse et d'évaluation, que nous sommes parvenus à concevoir l'architecture d'un environnement de développement multi-tactile et gestuel suffisamment ouvert pour répondre aux besoins de l'entreprise, aux contraintes soulevées par notre modèle d'interaction et aux résultats apportés par l'expérience utilisateur de nos premières applications.

C'est en 2010, au cours de la troisième année, que nous avons écrit notre environnement de développement, constitué d'un intergiciel multi-tactile gestuel en C++ et de deux interfaces de développement (API) pour *Adobe Flash (AS3)* et *Unity 3D (C#)*, des logiciels fréquemment utilisés par l'entreprise. Notre choix s'est porté sur ces derniers, non de manière exclusive puisque nous verrons que l'architecture système que nous avons conçue favorise le portage des API vers d'autres plateformes, mais pour répondre à une demande spécifique du groupe *DCNS* pour la réalisation d'une application à présenter sur le salon international *Euronaval* en octobre 2010. Nous avons alors défini notre cahier des charges et décidé des priorités de développement en fonction de cette échéance.

1. Extension du modèle d'interaction

Les premières applications que nous avons développées, lorsque le prototype est devenu opérationnel, concernaient principalement des exemples d'interfaces multi-tactiles sous la forme de maquettes fonctionnelles. Il s'agissait de pouvoir réaliser, devant les partenaires du projet et les clients de l'entreprise potentiellement intéressés, des démonstrations technologiques de l'écran multi-tactile. Dès les premières manipulations par des personnes extérieures à l'équipe de développement, et dans le cadre d'une procédure d'analyse des observations menées sur le terrain, nous avons remarqué que les utilisateurs éprouvaient parfois quelques difficultés à exécuter les tâches les plus simples comme déplacer un objet ou agrandir une image par pincement (*pinch*), des manipulations que certains connaissaient déjà pour les réaliser quotidiennement sur leur smartphone. D'autres ne parvenaient pas à identifier le type d'interaction qui était attendu par le système et manifestaient rapidement des signes d'agacement ou d'abandon.

Dans un premier temps, nous avons rejeté le problème sur le prototype, en pensant qu'il pouvait s'agir d'un problème d'illumination infrarouge, de sensibilité ou de calibration. Cependant, le fait que nous, concepteurs et développeurs du projet, n'étions que peu touchés par les difficultés ressenties par les utilisateurs, nous avons cru que ces problèmes pouvaient avoir des origines multiples et dépendaient non seulement du matériel mais également des environnements de développements opensources que nous utilisions et du manque de « *visibilité* » des interfaces que nous développons. Devant la fréquence élevée du phénomène, nous avons décidé de mettre en place et de généraliser une méthode d'analyse du comportement gestuel des utilisateurs par le biais d'enregistrements logiciels. De plus, nous avons profité de la participation de *Virtual-IT* à des salons ouverts au public pour réaliser ces enregistrements à travers de multiples projets et interfaces.

L'analyse des données tactiles nous a permis de mettre en évidence des facteurs que nous avions négligés dans le modèle conceptuel d'interaction ou dont nous n'avions pas conscience puisqu'ils étaient étroitement liés au matériel que nous utilisions. L'application pratique du modèle d'interaction sous la forme d'applications de démonstration s'est révélée incroyablement pertinente puisqu'elle nous a permis de valider et de préciser les hypothèses que nous avons formulées de manière théorique. De plus, à partir de ces

enregistrements, nous avons relevé des informations précieuses qui ont largement contribué à la conception de l'architecture de l'environnement de développement multi-tactile.

1.1. Méthodologie de recueil et de traitement des données

Matériel

Bien que nous ayons adapté le prototype pour le rendre démontable et moins encombrant, le coût de déplacement d'un tel dispositif et la configuration des espaces de démonstration lors des salons nous ont conduits à privilégier l'utilisation de surfaces plus petites comme une borne ou une table interactive. Pour construire ces dispositifs, nous avons utilisé une autre technique de captation des points de contact en utilisant des cadres optiques infrarouges. Ceci nous a permis de réduire l'encombrement causé par la projection vidéo et d'utiliser des écrans LCD et PLASMA pour gagner en luminosité, en contraste et en résolution.

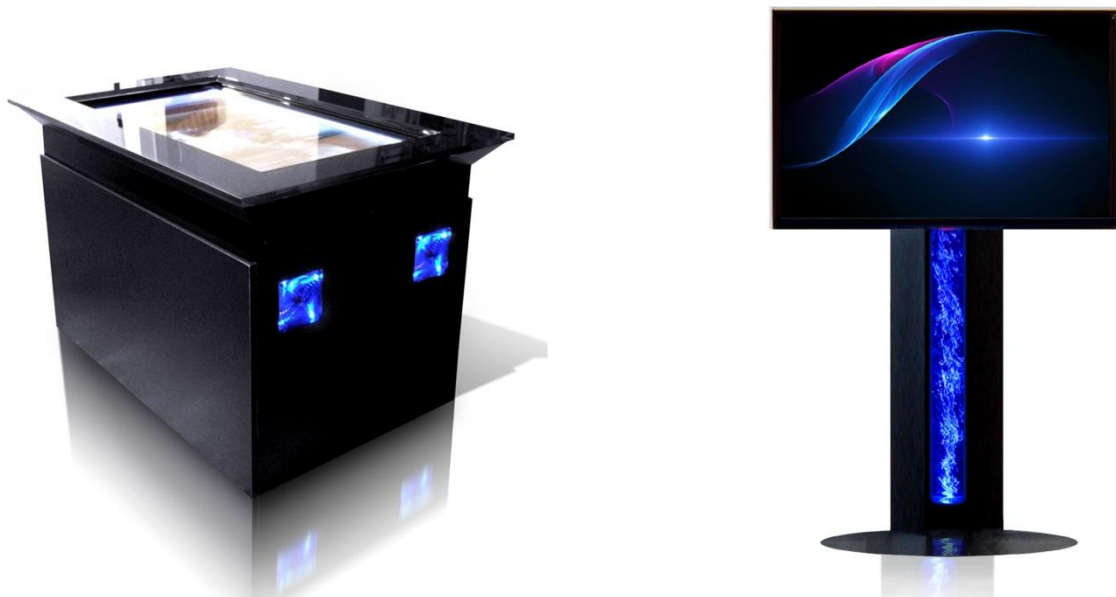


Figure 74 - A gauche, la table interactive. A droite, la borne multi-tactile.

Bien que la construction de ces dispositifs nous ait demandé un temps d'étude et de conception important, je ne m'attarderai pas sur les détails de leur fabrication car d'une part, ces informations sont soumises à la politique de confidentialité de l'entreprise et

d'autre part, elles n'apporteraient pas d'éclairage particulier sur le sujet qui nous intéresse. Toutefois, il me semble nécessaire de faire figurer ici les caractéristiques techniques du cadre optique infrarouge que nous avons utilisé lors de la phase d'enregistrement des données utilisateur :

- Taille de la surface : 50"
- Nombre de points de contact simultanés : 32
- Distance de détection du contact : quelques millimètres (variable selon la taille du doigt et les réglages effectués)
- Temps de réponse : 7 – 12ms
- Verre de protection : 3mm
- Points fantômes : aucun

Logiciel

La simple observation des utilisateurs en situation d'interaction ne suffisait pas toujours pour identifier l'origine des erreurs survenues lors des manipulations. Sur les interfaces que nous développons, aucun curseur n'apparaissait lors du contact des doigts sur l'écran car nous préférons utiliser des retours visuels personnalisés, en fonction des actions utilisateur reconnues par le système. De plus, nous ne voyions pas l'intérêt de doubler le retour haptique par une information visuelle de type curseur étant donné le faible degré d'indirection spatial de ces dispositifs qui rendait obsolète la représentation graphique de la position du pointeur en temps réel. D'autre part, nous craignons que le rappel visuel du curseur de la souris favorise chez l'utilisateur des interactions multiplexées dans le temps et non dans l'espace. Sans afficher la représentation des points de contact détectés par le système, il nous était difficile de comprendre l'apparition de certains comportements systèmes et les difficultés éprouvées par certains utilisateurs. C'est la raison qui nous a conduits à concevoir un système d'enregistrement des actions utilisateurs et des données tactiles afin de pouvoir analyser plus finement l'origine de ces phénomènes.

Ce système est composé de deux logiciels d'enregistrement :

- *CamStudio*, sous licence GPL, qui permet d'enregistrer les actions de l'utilisateur sur l'interface en capturant les images issues de l'application, en

temps réel, sous la forme de vidéo. Il est possible de spécifier le format, la cadence d'image et la zone précise de la capture.

- *TouchDataListener* (cf. figure 75), un logiciel développé à Virtual-IT, sous *Adobe Flex*, qui nous permet de visualiser l'enregistrement des données tactiles en affichant la représentation et les informations des points de contact et des événements tactiles.

Aujourd'hui, l'enregistrement des points dans le temps est effectué directement par l'intergiciel qui crée, dès la connexion d'une nouvelle application, un fichier texte au format CSV (*comma-separated values*) en lui donnant un nom unique grâce à un numéro de session, la date et l'heure. Dès la réception d'une donnée tactile, l'intergiciel convertit les informations en données exploitables par le logiciel *TouchDataListener*. Les messages sont enregistrés au format `TouchInfo` utilisé pour la communication entre l'intergiciel et l'API AS3 (*Adobe Flash*). Chaque événement `TouchInfo` est accompagné d'un temps en millisecondes qui permet d'enregistrer le moment exact de sa production.

```
<touchInfo>[TOUCH_TYPE]:[x]:[y]:[X]:[Y]:[A]:[dx]:[dy]</touchInfo>
```

Ces données sont ensuite traitées par l'application *TouchDataListener* qui effectue une analyse syntaxique (*parsing*) du fichier au format CSV et utilise les données de points pour proposer une représentation temporelle de l'activité tactile de l'utilisateur sous la forme d'un graphique (1) où l'on peut suivre les paramètres de vitesse et sous la forme d'un écran (2) qui affiche les points tactiles actifs et leur position. Ces deux représentations sont mises à jour automatiquement lorsque l'on déplace les poignets ou la barre de défilement (3) située sous le graphique. Grâce à ces contrôles, il est possible d'accéder rapidement à un moment précis du scénario, repéré sur l'enregistrement vidéo, et de réduire la zone temporelle à analyser pour gagner en détails. Enfin, en haut de l'interface, une zone (4) affiche l'état de la connexion entre l'intergiciel et l'application concernée au moment de l'enregistrement et reporte l'intégralité des messages bruts enregistrés dans le fichier CSV.

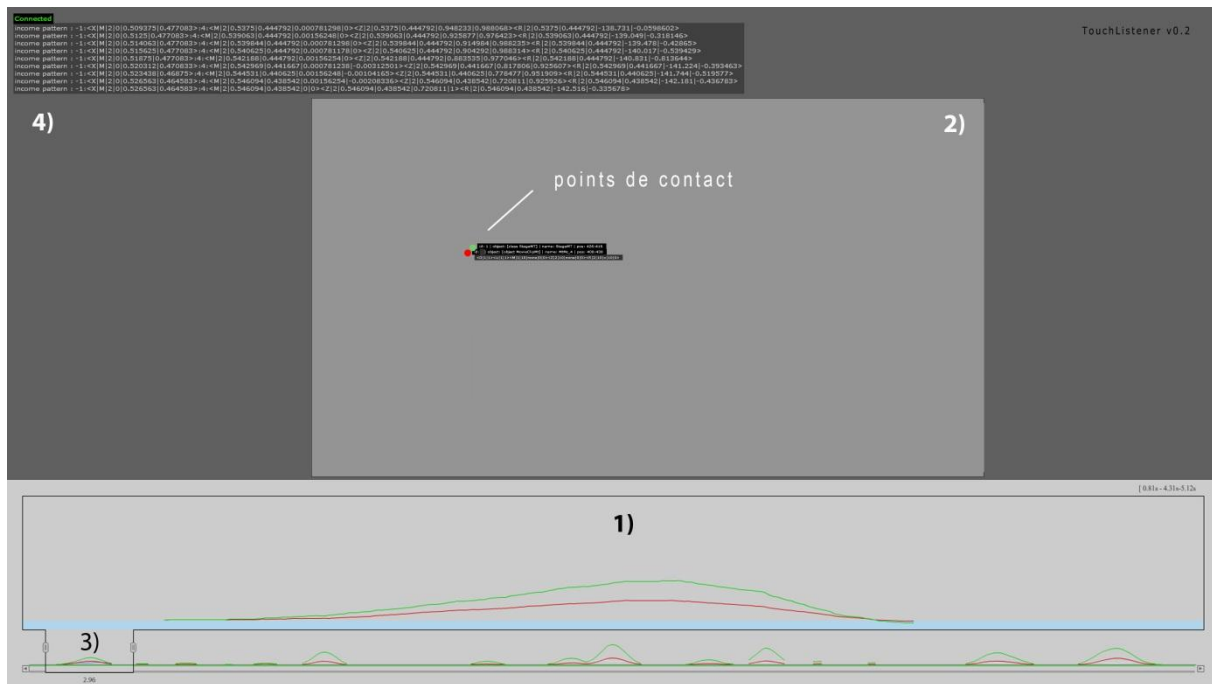


Figure 75 - TouchDataListener, outil de lecture d'enregistrements tactiles.

Pour nous permettre de retrouver facilement l'enregistrement correspondant à une observation réalisée sur le terrain, nous avons mis en place un système de prise de notes rapide où nous faisons apparaître les informations suivantes :

- La date et l'heure exacte du relevé
- L'application concernée
- La description du phénomène

Et lorsque cela était possible, nous tentions de répondre aux interrogations suivantes :

- Le phénomène s'est-il répété au cours de la session d'enregistrement ?
- La personne a-t-elle reçue des indications préalables sur les modalités d'interaction ?
- La personne a-t-elle déjà utilisé une interface multi-tactile ? Si oui, laquelle ?
- La personne reproduit-elle des gestes observés lors de sessions utilisateur précédentes ?

L'ensemble des enregistrements que nous avons en date du 11/08/2010 représente 116 sessions (une session par logiciel et par utilisateur), pour un temps d'enregistrement maximum de 2 minutes par session pour faciliter la lecture des résultats. D'une part, ces enregistrements nous ont permis de mettre en évidence l'origine de certains problèmes

d'interaction relevés sur site, après avoir noté l'application, la date et l'heure exacte de leur apparition. D'autre part, nous avons pu analyser les phrases gestuelles et relever des données précieuses pour la conception de l'environnement de développement.

1.2. Recueil des données

Dans ce court chapitre, je ferai l'inventaire des applications utilisées pour réaliser les enregistrements utilisateur, en détaillant le lieu et les interactions exploitées à travers chaque interface. Ces démonstrations ont eu lieu dans les locaux de Virtual-IT et dans des espaces public où les utilisateurs avaient libre accès au dispositif tactile. Les démonstrations ont été réalisées en essayant de couvrir un ensemble de configurations matérielles et logicielles différentes. Nous avons développé ces applications afin de faire évoluer notre modèle d'interaction, mais également pour répondre aux besoins de communication de l'entreprise.

La liste suivante n'est pas exhaustive mais correspond aux réalisations dont j'ai eu la responsabilité, en tant que chef de projet, et dans lesquelles je me suis également investi comme développeur ou intégrateur. Ces projets sont actuellement visibles dans le showroom de la société Virtual-IT et reflètent les travaux de recherche que nous avons menés au cours des deux premières années de thèse. Nous sommes parvenus, grâce aux efforts humains et financiers consentis par l'entreprise, à constituer un terrain d'expérimentation idéal pour l'évaluation de notre modèle d'interaction et la conception de l'environnement de développement multi-tactile. Le développement de ces projets internes a nécessité du temps et une organisation conséquente qui a impliqué le concours de l'ensemble des équipes de production de Virtual-IT.

Andromède

Lieu de l'enregistrement	:	Virtual-IT
Type de l'application	:	Navigateur / 3D
Technologie logicielle	:	<i>Nova (Vertice)</i>
Support	:	Table interactive et borne
Technologie tactile	:	Cadre optique (<= 2points de contact)
Utilisateurs	:	Professionnels de l'aménagement du territoire
Modalités d'interaction	:	1 ou 2 main(s) / simple utilisateur
Types d'interaction	:	

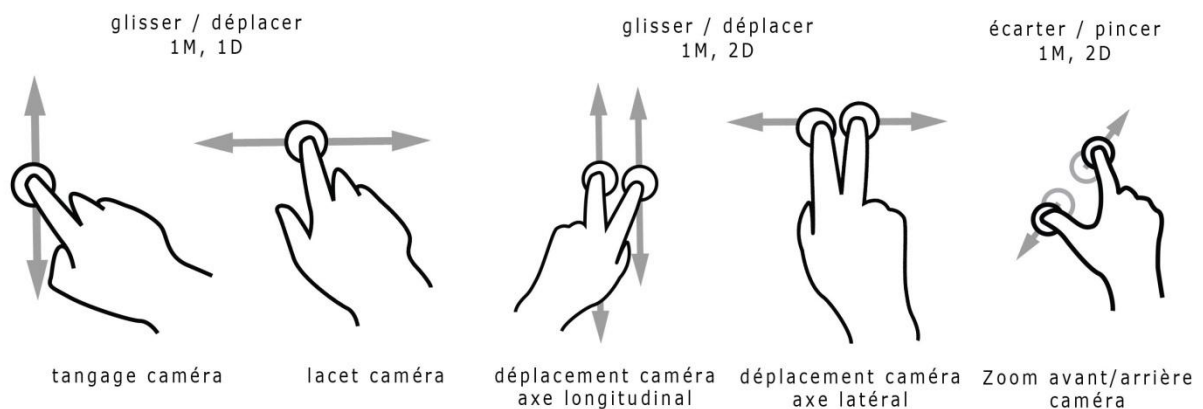


Figure 76 - Types d'interaction utilisés dans le projet.



Figure 77 - Capture d'écran du projet *Andromède*.

TouchMap

Lieu de l'enregistrement	:	Virtual-IT, INFOSUP
Type de l'application	:	Cartographie / 2D
Technologie logicielle	:	<i>Flash (Adobe)</i>
Support	:	Table interactive et borne
Technologie tactile	:	Cadre optique (<= 2points de contact)
Utilisateurs	:	Tout public
Modalités d'interaction	:	1 ou 2 main(s) / simple utilisateur
Types d'interaction	:	

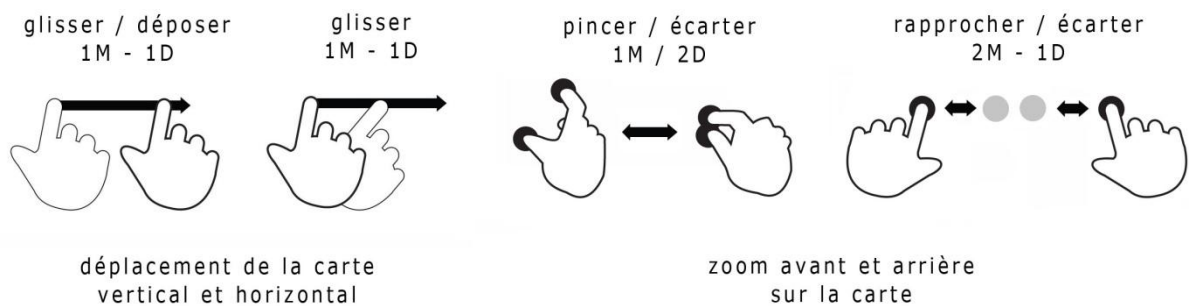


Figure 78 - Types d'interaction utilisés dans le projet.

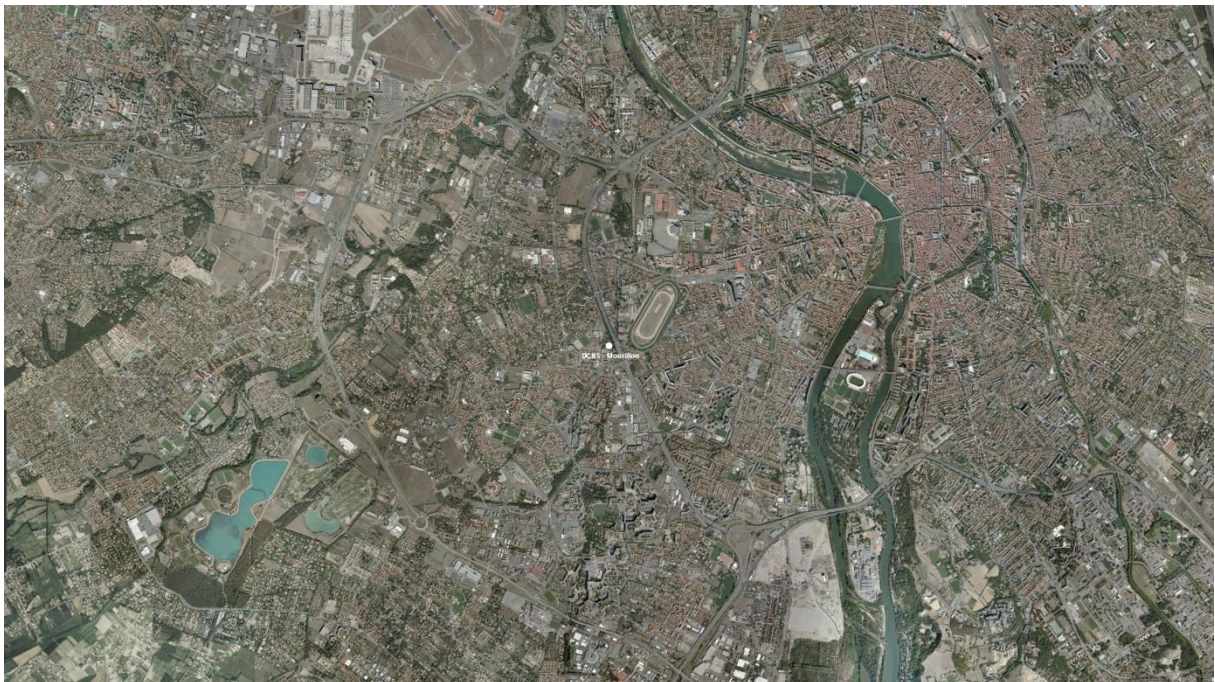


Figure 79 - Capture d'écran du projet *TouchMap*.

SPhIre

Lieu de l'enregistrement	:	Virtual-IT, IUT SéRéCom Castres, salon ARICAD
Type de l'application	:	Lecteur de médias / 3D
Technologie logicielle	:	<i>Flash (Adobe)</i>
Support	:	Table interactive et borne
Technologie tactile	:	Cadre optique (<= 2points de contact)
Utilisateurs	:	Tout public
Modalités d'interaction	:	1 ou 2 main(s) / simple utilisateur
Types d'interaction	:	

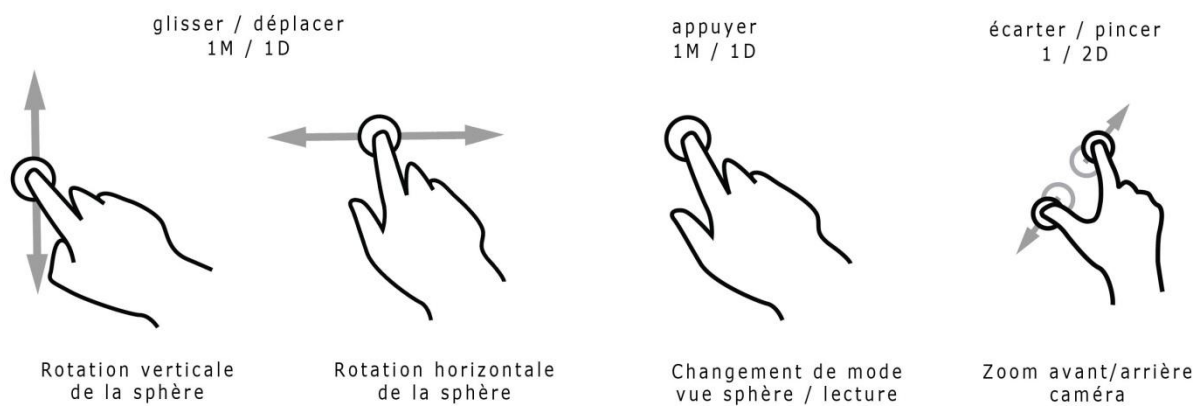
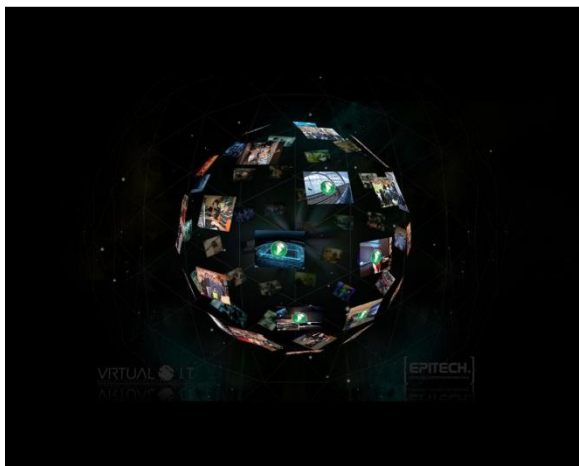


Figure 80 - Types d'interaction utilisés dans le projet.



vue sphère



vue lecture

Figure 81 - Capture d'écran du projet SPhIre.

Interface « carrousel »

Lieu de l'enregistrement	:	Virtual-IT, salon ARICAD
Type de l'application	:	Interface avec menu / 2D
Technologie logicielle	:	Flash (Adobe)
Support	:	Mur interactif, table, borne
Technologie tactile	:	DI (points illimités) , cadre optique (6 points)
Utilisateurs	:	Tout public
Modalités d'interaction	:	1 ou 2 main(s) / simple utilisateur
Types d'interaction	:	

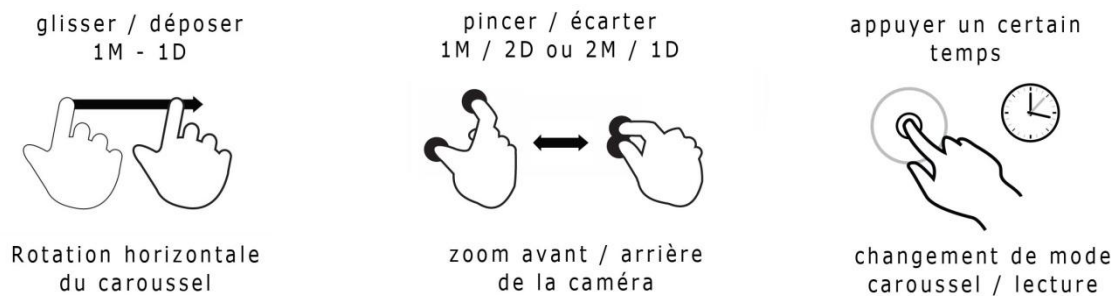


Figure 82 - Types d'interaction utilisés dans le projet.

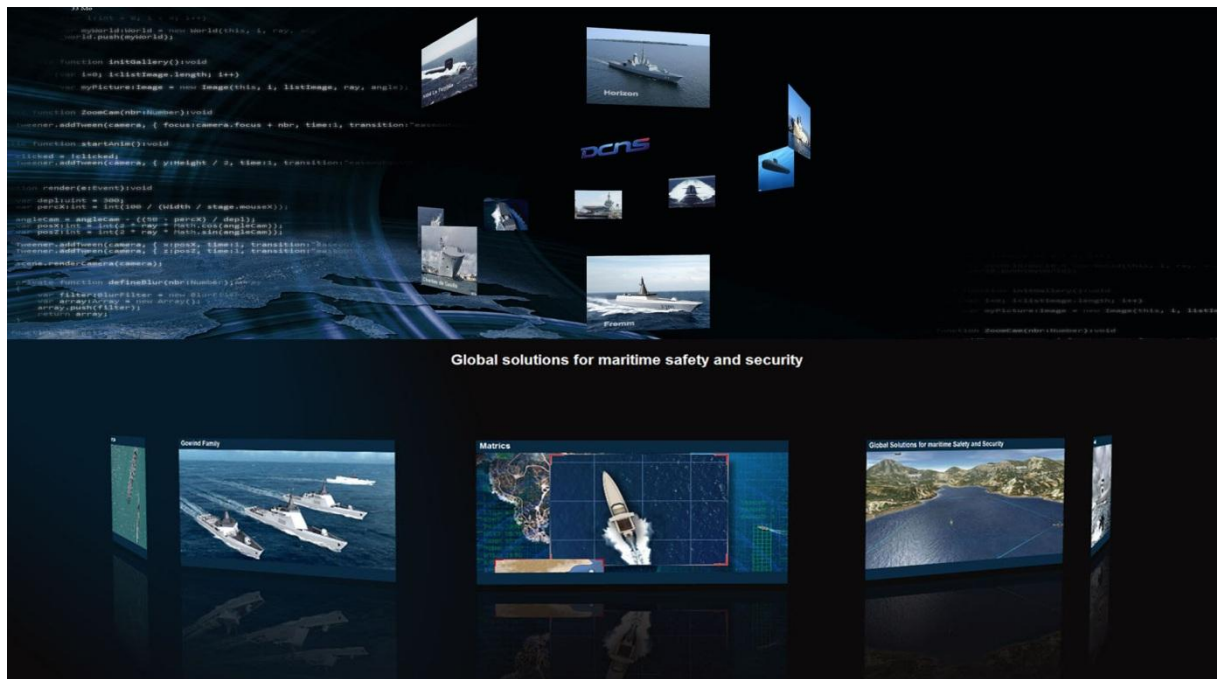


Figure 83 - Capture d'écran de deux projets avec un menu de type carrousel.

LUNAR

Lieu de l'enregistrement	:	Salon TGS (Toulouse Game Show)
Type de l'application	:	Navigation / 3D
Technologie logicielle	:	<i>Blender</i>
Support	:	Mur interactif
Technologie tactile	:	DI (points illimités)
Utilisateurs	:	Tout public
Modalités d'interaction	:	1 à 2 main(s) / multi-utilisateur
Types d'interaction	:	

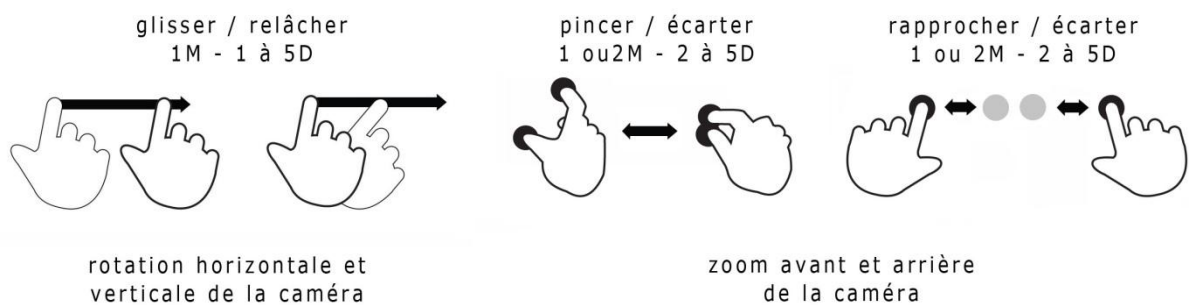


Figure 84 - Types d'interaction utilisés dans le projet.

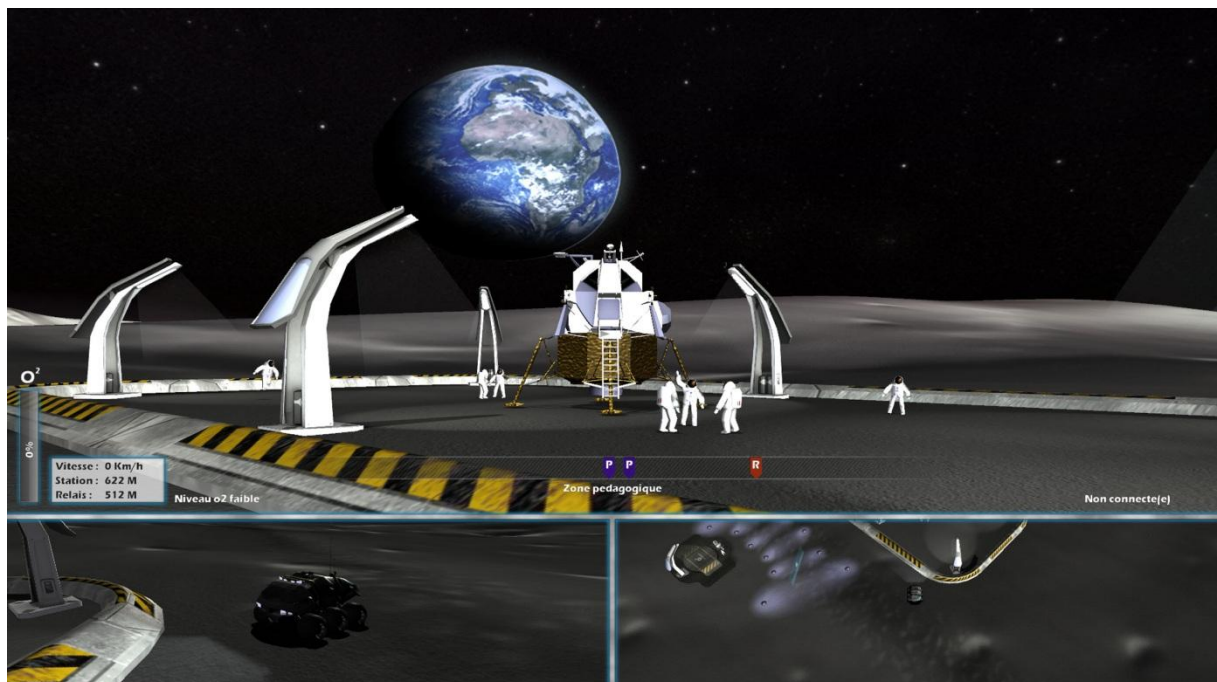


Figure 85 - Capture d'écran du projet LUNAR (trois vues utilisateur).

GSMSS (mur)

Lieu de l'enregistrement	:	DCNS (Toulon)
Type de l'application	:	Navigation / 3D
Technologie logicielle	:	<i>Blender</i>
Support	:	Mur interactif
Technologie tactile	:	DI (points illimités)
Utilisateurs	:	Responsables commerciaux et R&D de DCNS
Modalités d'interaction	:	1 à 2 main(s) / multi-utilisateur
Types d'interaction	:	



Figure 86 - Types d'interaction utilisés dans le projet.



Figure 87 - Capture d'écran du projet GSMSS (trois vues utilisateur).

GSMSS (table)

Lieu de l'enregistrement	:	DCNS (Toulon)
Type de l'application	:	Navigation / 3D
Technologie logicielle	:	FLASH
Support	:	Table
Technologie tactile	:	Cadre optique (2 points de contact)
Utilisateurs	:	Responsables commerciaux et R&D de DCNS
Modalités d'interaction	:	1 à 2 main(s) / simple utilisateur
Types d'interaction	:	



Figure 88 - Types d'interaction utilisés dans le projet.



Figure 89 - Capture d'écran du projet GSMSS version table, connecté en réseau à l'application GSMSS principale.

1.3. Ambiguïté d'échelle

La taille des écrans peut varier considérablement en fonction du domaine d'application. Ainsi un écran tactile multi-points peut mesurer 3,75cm par 4,09cm dans le cas d'un iPod nano comme il peut avoir une diagonale de plusieurs mètres dans le cas d'un écran large. Lors de la définition de notre modèle d'interaction, nous avons émis l'hypothèse que la taille de la zone interactive influençait considérablement le type d'interactions de l'utilisateur. Cependant, dans la pratique, nous avons observé des comportements gestuels qui nous ont permis de fixer les limites de ce postulat.

En effet, sur le terrain, nous avons remarqué un phénomène curieux en observant les utilisateurs interagir avec le mur interactif. En expliquant à l'utilisateur qu'il pouvait effectuer un zoom avant ou arrière sur un objet graphique (image, plan cartographique, etc.), celui-ci ne profitait pas de l'espace d'interaction qui lui était proposé mais réalisait le même geste que s'il agissait sur une petite surface. Bien que l'affordance perçue de l'objet à agrandir suggérait une interaction bi-manuelle et que l'utilisateur était libre de tout mouvement, celui-ci manipulait l'objet d'intérêt de grande dimension (>1m) avec un geste d'écartement du pouce et de l'index, sur quelques centimètres.

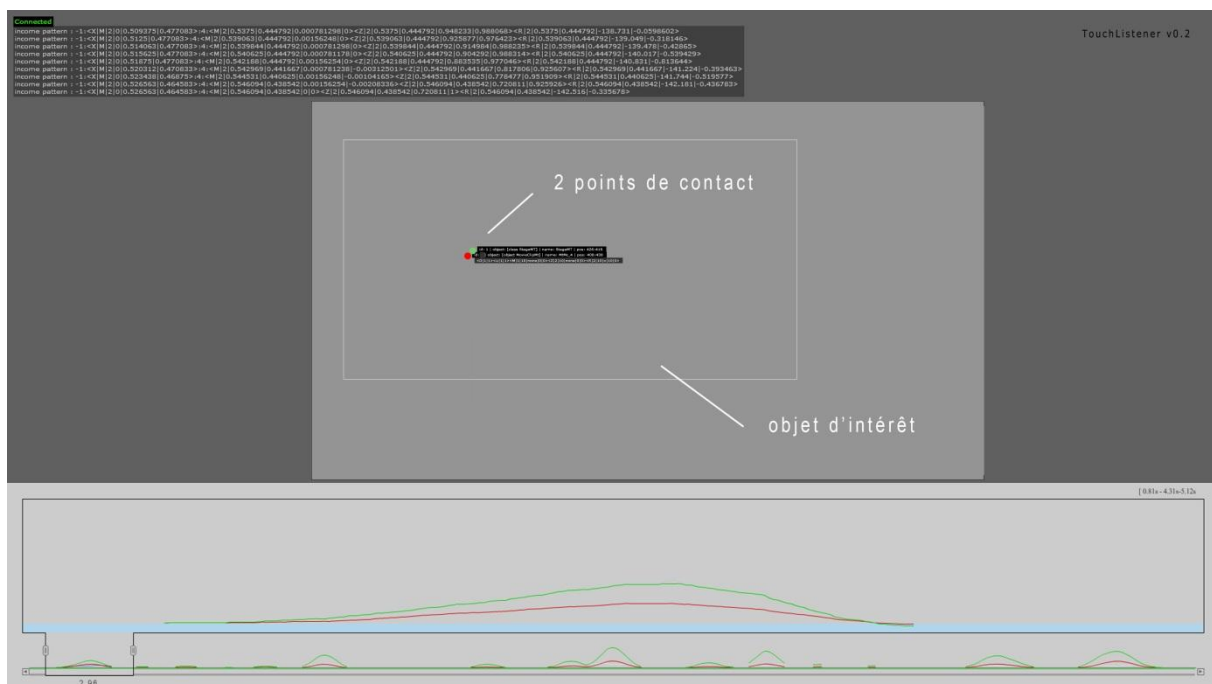


Figure 90 - Différence d'échelle observée entre l'interaction et l'objet d'intérêt. Composition créée à partir d'une capture d'écran de l'outil de lecture d'enregistrements *TouchDataListener*.

Pour calculer l'agrandissement relatif d'un objet graphique o en fonction de l'écartement des deux points les plus éloignés A et B , nous utilisons le rapport entre la distance de ces points en début d'interaction et la distance de ces mêmes points au moment du calcul. La différence d'écart initiale et actuelle entre le pouce et l'index était reportée sur la taille de l'objet. Ayant déterminé ce coefficient α , une simple règle de proportionnalité nous permettait de calculer la nouvelle dimension de l'objet en fonction de sa taille initiale. Un seuil minimum était fixé pour la distance initiale afin de ne pas obtenir un coefficient trop élevé.

$$\alpha = \frac{\sqrt{(x_{A'} - x_{B'})^2 + (y_{A'} - y_{B'})^2}}{\sqrt{(x_A - x_B)^2 + (y_A - y_B)^2}}$$

Etant donné la forte différence d'échelle entre les paramètres de la manipulation et ceux de la modification visuelle de l'objet, l'utilisateur n'effectuait pas l'interaction d'écartement avec suffisamment de précision pour avoir un contrôle efficace sur le système. En effet, écarter ses doigts de quelques millimètres sur la surface interactive revenait à agrandir l'image de plusieurs dizaines de centimètres. Plusieurs facteurs peuvent être à l'origine de cette ambiguïté spatiale. Premièrement, les utilisateurs, présents sur les salons, n'avaient pas toujours les deux mains libres. Rares étaient ceux qui posaient leurs affaires, ou les confiaient aux personnes les accompagnant, afin de bénéficier des interactions bi-manuelles suggérées par la taille du dispositif et par l'affordance perceptible des objets graphiques. Deuxièmement, certains zooms, avant ou arrière, étaient réalisés non pas sur des objets mais sur des containers, comme c'est le cas dans l'application *LUNAR* ou *GSMSS*. Nous pouvons supposer que la nature abstraite de cette manipulation et l'absence d'affordance perceptible d'une vue en trois dimensions ne suggèrent pas suffisamment les modalités d'une interaction bi-manuelle. Enfin, comme cela nous l'a été confié par certains utilisateurs, l'idée du geste de pincement qu'ils ont effectué sur l'objet d'intérêt, avec le pouce et l'index, leur est venue par le réflexe né de l'apprentissage des interactions multi-tactiles sur leur smartphone. Ce constat appuie l'hypothèse émise au cours de notre modèle d'interaction qu'une standardisation gestuelle trop précoce ou trop tardive peut avoir des influences considérables sur l'apprentissage de nouvelles modalités d'interaction.

1.4. Points aberrants

La première application que nous avons développée était un lecteur de médias (photos et vidéos) dont les éléments affichés à l'écran étaient sélectionnables et manipulables. Il était possible de déplacer un média avec un doigt ou de modifier sa taille et sa rotation avec deux doigts. Bien que nous ayons implémenté un moteur physique²⁰¹ aux éléments de l'interface, et donné une certaine inertie aux mouvements des objets graphiques, nous avons remarqué des comportements étranges de ces objets lors de leur manipulation. En effet, au cours de leur transformation ou de leur déplacement dans le temps, les propriétés de l'objet prenaient soudainement, et pour une courte durée (quelques images), des valeurs « aberrantes », comme si par moment, les données tactiles étaient erronées.

En analysant les données issues de l'enregistrement (cf figure 91) effectué sur la borne interactive, nous avons effectivement remarqué un pic de vitesse du deuxième point de contact (id=1) tout à fait anormal puisque selon le graphique (1), la vitesse aurait été multipliée par 3 d'une image à l'autre, soit en $1/60^{\text{ème}}$ de seconde. En vérifiant les positions des points de contact sur la représentation virtuelle de l'écran au moment du pic, nous avons fait le constat qu'un des points de contact, pour une durée de quelques images seulement, changeait subitement de position, comme si le signal était parasité.

En essayant de reproduire le phénomène, nous avons constaté qu'il était la conséquence d'un contact involontaire de la part de l'utilisateur. Il est possible, étant donné la sensibilité élevée des technologies tactiles que nous utilisons, que ce point de contact provienne de l'effleurement de la veste de l'utilisateur sur l'écran lors d'une manipulation. Cela expliquerait pourquoi nous n'avons pas encore observé ce problème en laboratoire. Par la suite, nous avons également fait le constat que certains utilisateurs, en touchant l'écran avec l'index pointé et les autres doigts repliés, pouvait également toucher l'écran au niveau de l'articulation interphalangienne distale du majeur (cf. figure 53), ou du bout des autres doigts en fonction de la configuration spatiale interne de la main par rapport à la surface.

²⁰¹ Projet open source *box2d*. <http://box2dflash.sourceforge.net/>

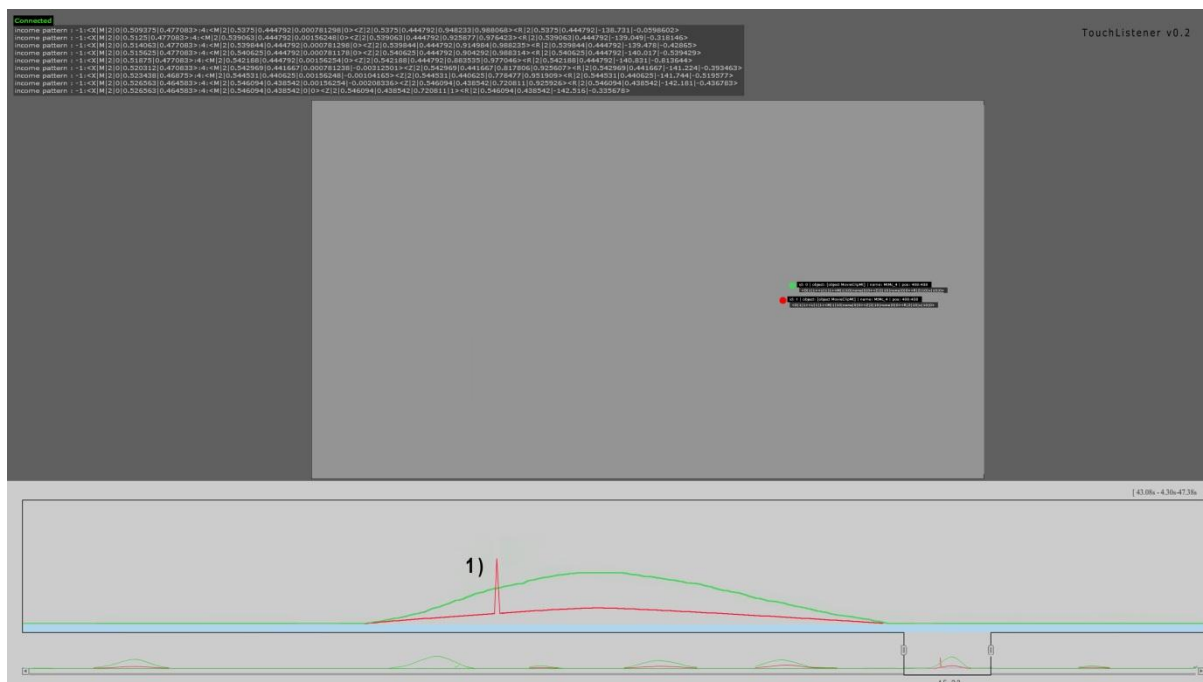


Figure 91 - Observation de points aberrants. Composition créée à partir d'une capture d'écran de l'outil de lecture d'enregistrements *TouchDataListener*.

Or, les cadres optiques que nous utilisons au début de notre recherche n'étaient capables de détecter que deux points de contacts simultanément. Lorsqu'un troisième était effectué, le microcontrôleur reportait une position erronée d'un des deux points existants, sans changer son identifiant, il était donc impossible pour nous de corriger correctement ce problème de manière logicielle. Nous avons tout d'abord sous-estimé ce problème, en pensant qu'il disparaîtrait rapidement avec l'évolution de la technologie. Toutefois, lorsque nous avons fait le portage de cette application vers le mur interactif, dont la technologie d'acquisition optique par caméra nous assurait de ne pas générer de reports erronés de positions, nous avons fait un nouveau constat : le nombre de points de contact pouvait changer au cours d'une même phrase gestuelle et créer des phénomènes similaires à ceux produits sur les cadres optiques. Bien que cela puisse nous paraître évident aujourd'hui, lors du développement de cette application, nous avons défini les interactions entre les points de contact et les objets graphiques selon le couplage suivant :

- Le déplacement d'un objet graphique est déterminé par la position du premier point de contact (point de référence d'ID 0) qui lui est associé et par son évolution dans le temps. Si des points de contact étaient ajoutés ou supprimés au cours du temps, ils étaient ignorés.

- L'agrandissement et la rotation d'un objet sont définis par les positions et les mouvements des deux points de référence (d'ID 0 et 1) les plus éloignés qui lui sont associés. Si des points de contact étaient ajoutés ou supprimés au cours du temps, ils étaient ignorés.

Or, ce principe ne corrige que partiellement les conséquences de l'introduction de points supplémentaires. Cela ne fonctionne que si les points de référence sont présents tout au long de la phrase gestuelle. En analysant les enregistrements effectués sur le mur interactif (43 sessions), nous avons relevé que près d'un tiers des phrases gestuelles, dont l'interaction implique de 1 à 5 doigts, subit une variation du nombre de points de contact au cours du temps et que dans près de la moitié des cas, cette variation entraîne des changements du ou des points de référence (cf. figure 92). Cette analyse s'est révélée déterminante puisque nous avons tenu compte de ces paramètres en implémentant un système de correction de points dans l'intergiciel (cf. Chapitre VI - 2.4.1.)

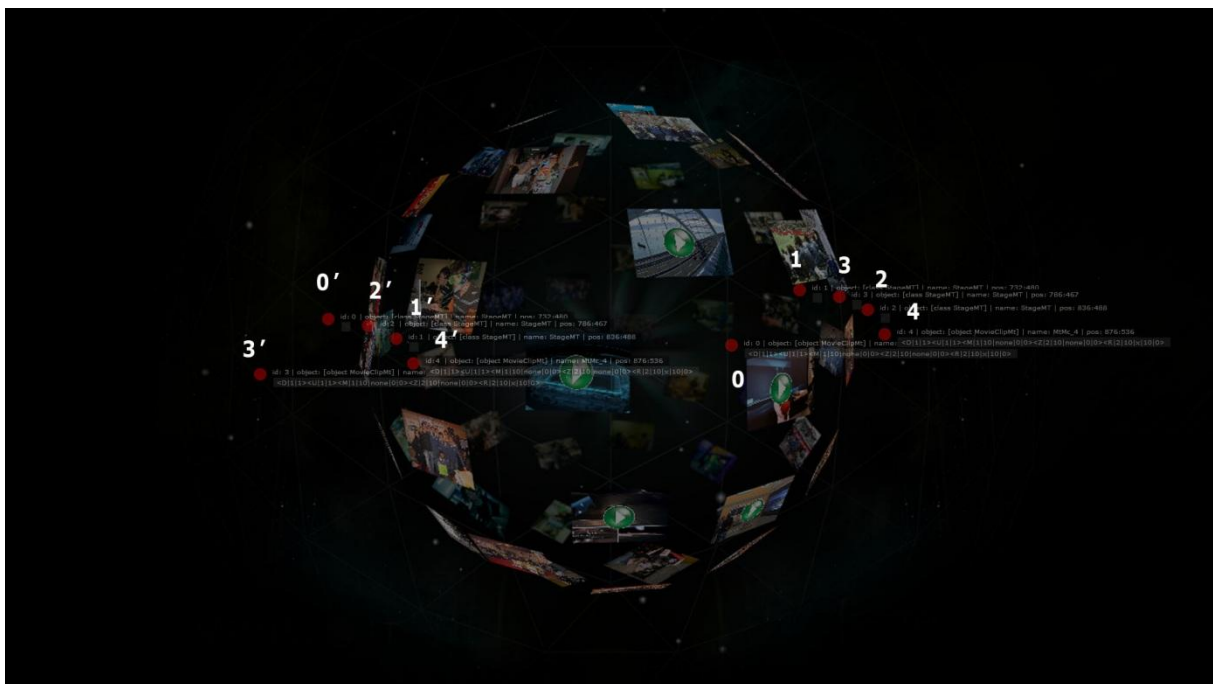


Figure 92 - Illustration du changement de l'ID des points de contact au cours de la phrase gestuelle. Composition créée à partir d'une capture d'écran du projet Sphire et des données issues de la lecture des enregistrements par l'outil *TouchDataListener*.

1.5. Sensibilité des systèmes optiques

Le cas de figure que nous avons relevé dans la partie précédente où, selon la technique de captation, des points de contact involontaires peuvent surgir, soulève la question plus

générale de la sensibilité des dispositifs que nous avons utilisés. Lors des phases d'observation des utilisateurs que nous avons rencontrés sur le terrain, il est apparu plusieurs cas où les interactions avec le système semblaient manquer de précision. En nous penchant sur le problème, nous avons mis en évidence trois moments de la phrase gestuelle où les intentions de l'utilisateur et la réaction du système n'étaient pas coordonnées et où le feedback visuel était en décalage par rapport aux réactions attendues :

La fin de la phase de préparation du geste. Il s'agit du moment où la main, en s'approchant de la surface interactive, se positionne avant d'effectuer les contacts d'interaction. À ce moment précis, l'utilisateur ne dispose pas encore de retour visuel pour guider son mouvement et n'a pas reçu de retour haptique lui confirmant qu'il a bien touché l'écran. Cependant, un problème de sensibilité commun à une grande majorité des technologies optiques donne au système la capacité de détecter un point de contact avant même le contact physique réel entre les doigts de l'utilisateur et l'écran. L'utilisateur n'en a pas conscience puisqu'aucune indication ne lui permet de savoir à quel moment précis le système va réagir. Ce constat nous invite à reconsidérer le postulat que nous avons émis précédemment qui supposait d'écarter, de manière théorique, la phase de préparation du geste puisque nous avons jugé que la nature des systèmes tactiles n'impliquait de s'intéresser qu'à la phase nucléaire de la phrase gestuelle.

Le début de la phase nucléaire. Il s'agit du moment où la ou les mains de l'utilisateur sont posées sur la surface interactive, et où l'utilisateur commence, de manière intentionnelle, le geste d'interaction ou de manipulation. À cet instant, le système n'a pas encore déterminé l'action de l'utilisateur. Le début de cette phase est souvent caractérisé par la position statique de la main et du bras qui correspond à la fin du geste de pointage des doigts qui se posent sur un objet précis de l'interface. Toutefois, la sensibilité de certaines technologies, les micro-mouvements non intentionnels de l'utilisateur et le manque de précision de certains gestes, conduisent le système à reconnaître des actions qui ne sont pas véritablement voulues. Ce phénomène intervient au début de la phase nucléaire, au moment où l'utilisateur a les doigts sur l'objet d'intérêt. Soit l'utilisateur a commencé un geste et dans ce cas le système n'a pas encore reconnu l'interaction en cours puisque l'amplitude du mouvement n'est pas suffisante. Soit la pause statique du geste est souhaitée par l'utilisateur (maintien d'une zone de texte avant défilement) mais il ne parvient pas à stabiliser l'objet d'intérêt ou la manipulation en cours à cause de la sensibilité des capteurs.

Le début de la phase de rétractation. De manière similaire aux problèmes soulevés au moment de la fin de la phase de préparation, le début de la phase de rétractation, qui intervient après le soulèvement des doigts de la surface, peut entraîner des contacts involontaires avec l'écran et générer des interactions non intentionnelles. Ce problème s'est manifesté notamment lors du déplacement d'un objet graphique auquel nous avons calculé l'inertie sur les derniers moments cinétiques du mouvement. Parfois, le contact non intentionnel d'une partie de la main lors de la rétractation a eu pour conséquence la surestimation de cette valeur et l'apparition de déplacements aberrants de l'objet d'intérêt.

Suite à ces observations, nous revenons sur le modèle d'interaction théorique, défini au chapitre IV, en apportant quelques précisions sur la modélisation spatiale du geste et les différents états du système. Dans un premier temps, il nous semble nécessaire de replacer les phases de préparation et de rétractation du geste dans la phrase gestuelle des systèmes multi-tactiles optiques. Les données issues de la capture du geste, avant et après sa phase nucléaire, peuvent être la source des problèmes d'interaction que nous avons cités mais elles pourraient, à condition d'être suffisamment riches et précises, profiter au design d'interaction. En connaissant par exemple l'angle d'approche de la main lors de la phase de préparation, le système pourrait anticiper les actions de l'utilisateur en mettant en évidence (zoom avant) l'objet visé par le geste déictique. Cette considération nous a en partie amené l'idée que nous abordons en conclusion, sur les perspectives d'évolution de l'intergiciel et la conception d'un système d'information didactique et temps réel des interactions gestuelles.

Comprendre ces phénomènes nous a également permis d'apporter des solutions concrètes, dans le développement de l'intergiciel multi-tactile, aux problèmes liés à la sensibilité des technologies optiques. D'une part, nous avons implémenté des algorithmes de lissage de mouvement (cf. chapitre 6 - 2.4.2.) afin de limiter les effets des points aberrants, des mouvements saccadés ou non intentionnels de l'utilisateur. D'autre part, nous avons implémenté un principe de reconnaissance des interactions par dépassement d'un seuil de confiance (delta d'activation) afin d'éviter que le système, dans le cadre d'interactions continues (non discrètes), ne soit trop influencé par l'imprécision des mouvements réalisés par l'utilisateur.

2. Développement de l'intergiciel multi-tactile

2.1. Architecture logicielle

Limitations de l'existant

Lors de nos premiers développements sur plateforme multi-tactile, nous avons eu recours à des interfaces de programmation (API) existantes, soit fournies par le constructeur des cadres optiques que nous utilisons, soit mises à disposition par des développeurs indépendants sous licence libre. Les cadres optiques étaient livrés soit avec un pilote compatible HID (*Human Interface Device*), soit avec un serveur de communication des données. La liaison entre l'application et le cadre optique s'effectuait alors par l'intermédiaire d'une DLL ou d'un client socket. En ce qui concerne les technologies compatibles TUIO, nous utilisons des passerelles TUIO/XML Socket pour le développement d'applications Adobe Flash, ou nous implémentons des libraires libres pour connecter nos applications C++ (ogre3D), Python (Blender), C# (Unity 3D). Enfin, les API nous permettaient d'accélérer les temps de production des prototypes logiciels que nous développions.

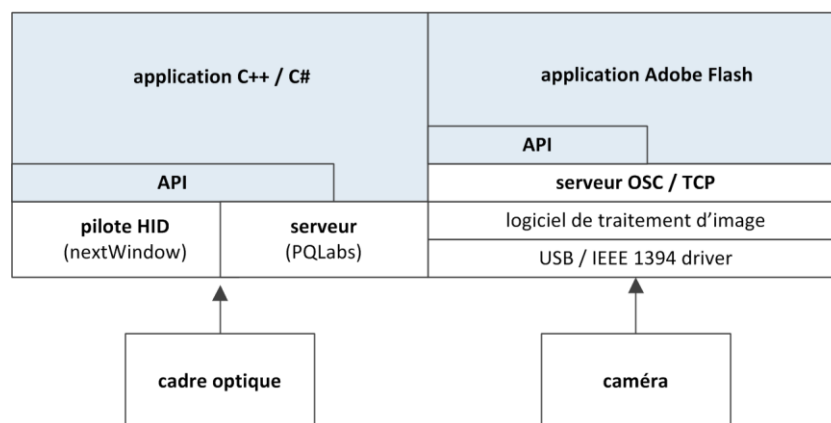


Figure 93 - bloc-diagramme des architectures logicielles multi-tactiles existantes pour les applications C++ / C# et Adobe Flash.

Cependant, lorsque nous devons présenter en public les projets sur lesquels nous travaillons, nous nous sommes retrouvés confrontés à des problèmes évidents de compatibilité entre les applications et les dispositifs multi-tactiles.

D'une part, quand nous avons dû faire le portage de certaines applications, d'une technologie tactile à une autre, nous y avons passé beaucoup de temps car il existait des

différences considérables dans la gestion des événements tactiles et dans les logiques de programmation d'une API à l'autre. D'autre part, l'évolution des technologies de captation optique, la mise à jour des micro-logiciels (*firmwares*) et celle des pilotes et API correspondants, ont considérablement porté préjudice à la pérennité des applications que nous avons développées. Enfin, nous étions assez fréquemment contraints par les API, soit car elle n'existait pas pour une plateforme logicielle donnée et dans ce cas, nous devions développer entièrement l'API (communication serveur, algorithmes d'analyse et d'interprétation des gestes, manipulation de la couche graphique, etc). Soit car la manière dont elles avaient été conçues ne convenait pas à nos besoins et nous devions reprogrammer une grande partie des classes pour dépasser les limitations de ces interfaces de programmation.

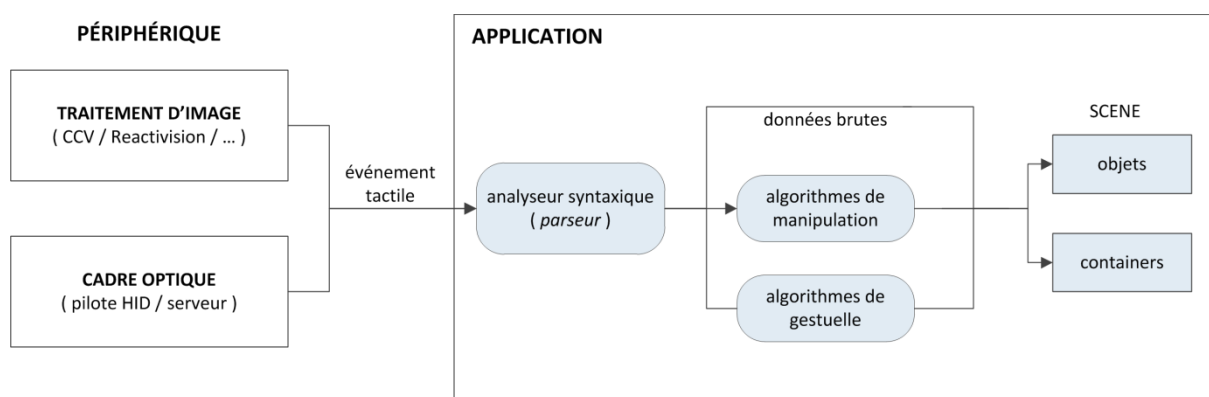


Figure 94 - Organigramme fonctionnel des architectures logicielles multi-tactiles existantes.

Proposition

C'est en faisant le constat que les API d'origine nous imposaient davantage de contraintes qu'elles ne nous offraient d'avantages, que nous avons décidé de prendre le temps de développer notre propre environnement de développement (*framework*). La décision a été délicate à prendre pour l'entreprise, partenaire du CIFRE, car ce choix allait convoquer le travail de trois personnes sur plusieurs mois. En considérant que je travaillais à temps partiel et que les deux développeurs de l'équipe que j'encadrais travaillaient en même temps sur d'autres projets, nous avons estimé un temps de développement de six mois pour la conception et le développement d'un environnement de développement multi-tactile gestuel. Malgré l'investissement financier et le temps qu'allait engager ce développement, l'entreprise a reconnu l'intérêt que nous avions à travailler dans ce sens. Pour répondre aux problèmes soulevés précédemment, et dans l'optique d'unifier les développements, passés

et futurs, nous avons décidé de développer notre propre environnement de développement multi-tactile gestuel, composé d'un intergiciel et d'interfaces de programmation pour les différentes plateformes logicielles utilisées par l'entreprise. En considérant le peu de temps qu'il restait avant la fin du contrat CIFRE, nous avons choisi de créer deux API, en AS3 pour les applications Flash et Flex Builder, et en C# pour les applications développées sous le logiciel 3D temps réel : Unity.

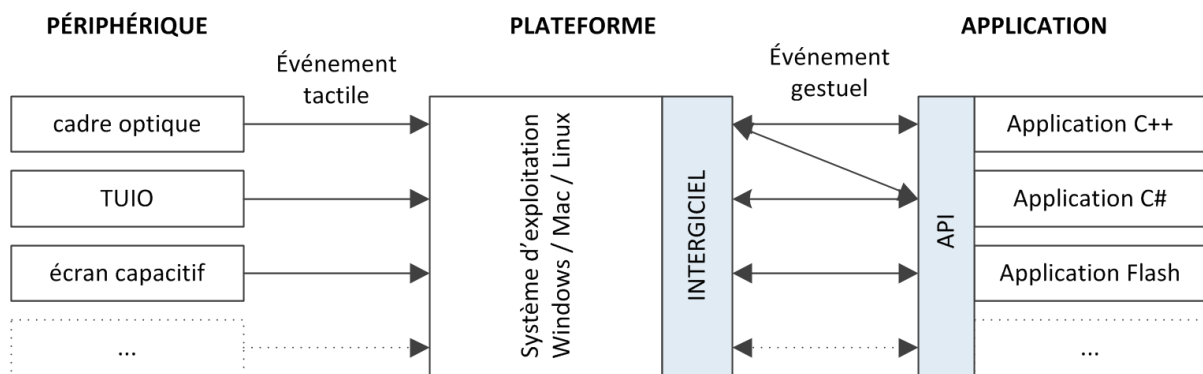


Figure 95 - Organigramme simplifié de l'architecture de l'intergiciel.

L'environnement de développement que nous avons élaboré répondait à plusieurs objectifs. D'une part, nous souhaitions unifier le développement d'applications multi-tactiles en dissociant l'API de la technologie tactile. L'idée était de pouvoir faire des développements multiplateformes, tant sur un plan logiciel (compatibles avec plusieurs systèmes d'exploitations) que matériel (compatibles avec plusieurs technologies tactiles). C'est pourquoi nous avons privilégié l'utilisation du langage C++ et la librairie graphique QT pour assurer la recompilation du code pour d'autres environnements comme Unix et Mac OS. De plus, nous nous sommes affranchis de la dépendance de l'application au périphérique d'entrée multi-tactile, en créant un format unifié et générique de description de points de contact. Lorsqu'une technologie évolue, au lieu d'être obligé de modifier l'ensemble des applications déjà réalisées, une simple mise à jour de l'intergiciel, par l'ajout d'une nouvelle librairie de périphérique, suffit pour assurer la compatibilité des applications.

D'autre part, nous souhaitons simplifier le développement et réduire les temps de production nécessaires au développement d'une application. Déplacer conceptuellement le traitement et l'analyse des points de contact, la reconnaissance gestuelle et les algorithmes de lissage de l'application vers l'intergiciel, nous a permis d'alléger le code à produire côté

application et de réduire le temps de développement d'une API pour une nouvelle plateforme logicielle (langage de programmation, environnement de développement, etc.)

Pour parvenir à ce résultat, nous avons élaboré une architecture logicielle originale puisque contrairement au fonctionnement traditionnel des API multi-tactiles (cf. figure 94), tous les traitements de données tactiles et gestuelles sont effectués par l'intergiciel. Pour tout nouveau point de contact, l'intergiciel se charge de l'associer à l'objet graphique ou au container situé dessous en interrogeant l'application active (située au premier plan) et en récupérant l'intégralité des propriétés et des comportements tactiles de l'objet. L'idée est de déplacer l'ensemble des processus d'analyse et de traitement des points de contact de l'application vers l'intergiciel et d'unifier ainsi les développements. La définition de ces propriétés concerne ce que nous appelons les patrons (*patterns*) d'interaction (cf. chapitre VI - 2.3.4.), que nous avons classés sous deux groupes : les interactions de manipulation et les interactions gestuelles.

L'intergiciel peut gérer plusieurs applications de manière simultanée, puisque chaque application, lorsqu'elle devient active en passant au premier plan, envoie à l'intergiciel un identifiant de session unique et propre à chaque application. Envoyée par connexion TCP, ces données, que nous avons nommées THA pour « *Touch Header Application* », sont définies par la syntaxe de description balisée suivante :

```
<AppName>[type d'application]:[numéro de session]</AppName>
```

Cadre de développement

Pour mener à bien ce projet, nous avons choisi le langage C++ en prévision d'un portage sur d'autres plateformes et pour sa rapidité d'exécution. En effet, comme nous l'avons soulevé précédemment, la latence du système est une contrainte particulièrement déterminante pour la qualité des feedbacks impliqués dans les interfaces tactiles.

Comme la majorité des applications développées par l'entreprise fonctionne sous environnement *Microsoft Windows*, c'est la plateforme que nous avons privilégiée pour les premiers développements.

Le temps de développement de l'ensemble de l'environnement de développement, intergiciel (version *Windows*) et API (*Adobe Flash* et *Unity*), ainsi que deux applications de

démonstration présentées lors du salon international *Euronaval* 2010, estimé initialement sur six mois, s'est finalement étalé sur une période de huit mois. Les trois années de contrat CIFRE se sont conclues par l'aboutissement d'une première version, déboguée et fonctionnelle de l'environnement de développement.

Le diagramme de classe (cf. figure 97) présente les différentes classes qui composent l'intergiciel. Ces classes sont regroupées dans les packages suivants :

`Devices` concerne les périphériques d'entrée et la conversion des touch en point générique.

`IDAI` est dédiée à la gestion des interactions tactiles.

`Core` est le noyau de l'intergiciel. Il contient la classe `SMT` qui instancie le périphérique d'entrée (device) ainsi que la classe `Sor`. La classe `Sor`, quant à elle, assure l'interface entre les données converties et l'application cliente par l'intermédiaire de l'`IDAI`.

`Patterns` contient les classes gérant les interactions tactiles et gestuelles.

`Legacy` concerne la rétro-compatibilité avec les applications utilisant `TUIO` et `XMLSocket`.

`Socket` contient les classes assurant les connexions clients/serveurs.

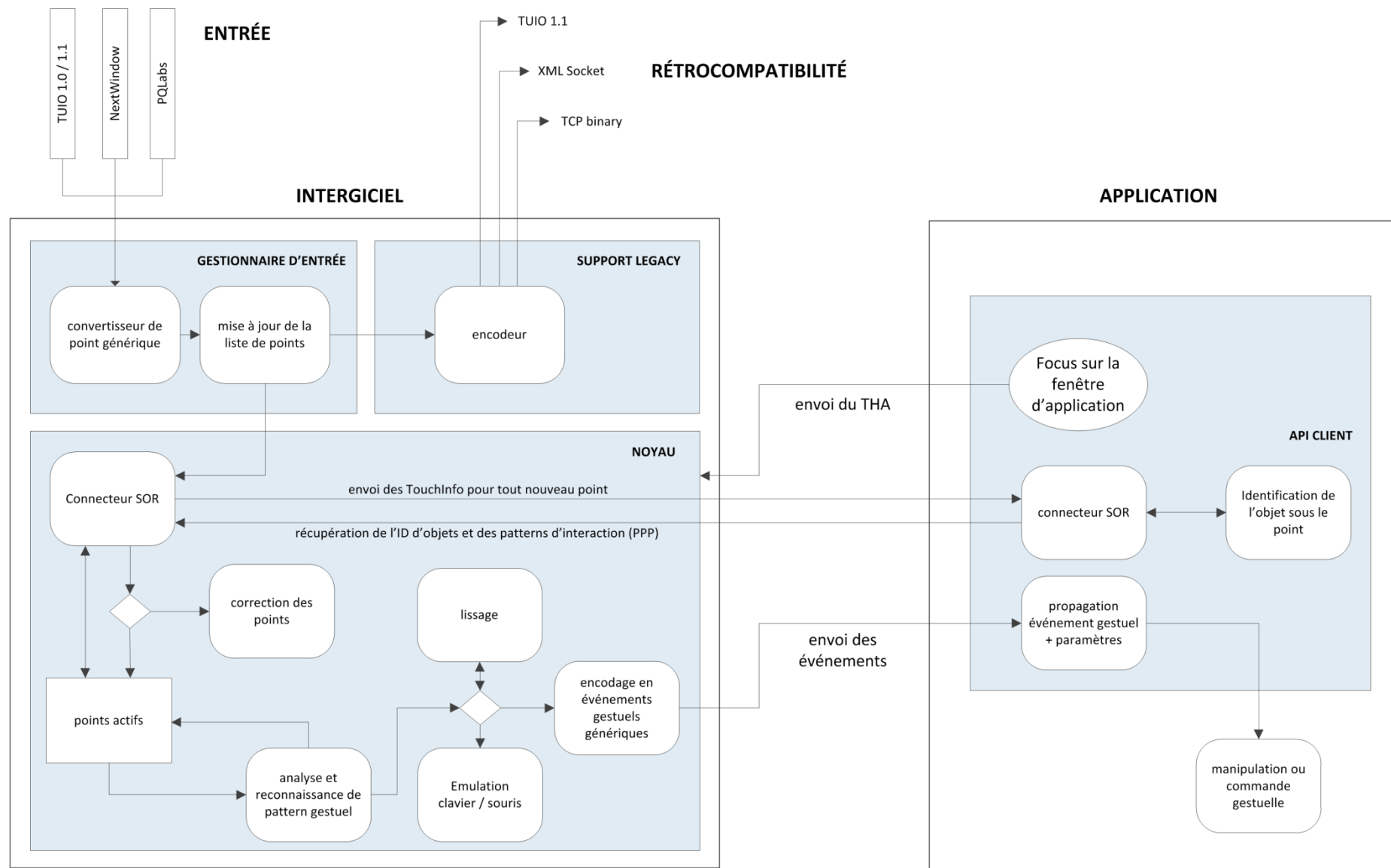


Figure 96 - Organigramme fonctionnel de l'environnement de développement (en bleu). Nous avons conçu une architecture qui unifie la définition de points de contact et leur analyse. Enfin, notre modèle peut être étendu à de nouveaux types d'interaction, de nouvelles plateformes et de nouveaux langages de programmation.

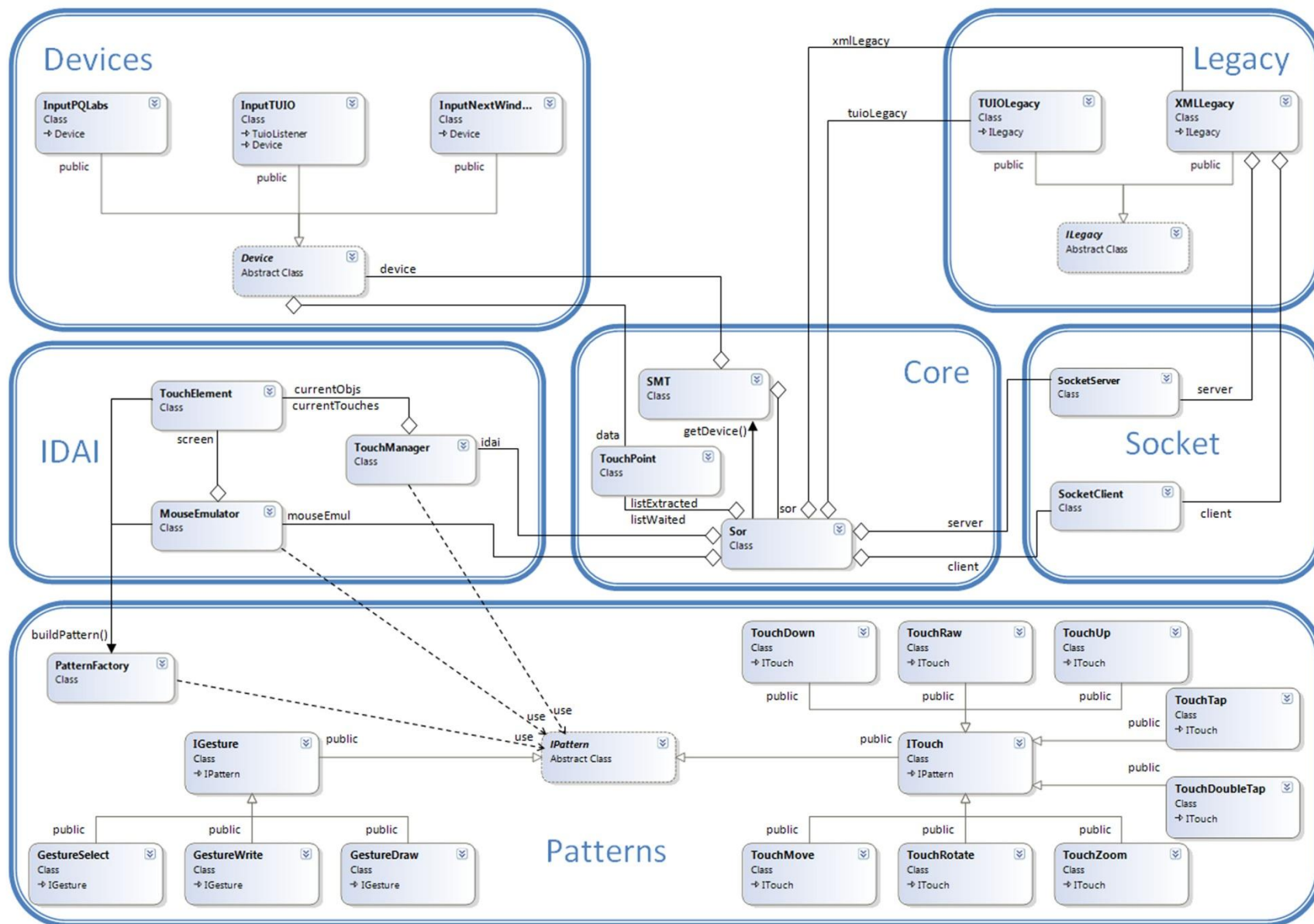


Figure 97 - Diagramme de classe de l'intergiel.

2.2. Données tactiles

Récupération des données tactiles

Pour assurer la récupération et la conversion des données tactiles issues des trois périphériques (*PQLabs*, *NextWindow* et *CCV*) sélectionnés pour la première version de l'intergiciel, nous avons mis en place un système de plugin, qui peut être alimenté par de nouvelles librairies que nous développerons dès qu'un périphérique change de version ou dès qu'une nouvelle technologie d'acquisition est intégrée aux dispositifs commercialisés par l'entreprise. L'intergiciel identifie automatiquement, en vérifiant l'état de la connexion pour chaque périphérique, quel est celui qui est actuellement relié au système. La méthode `extract()` de la classe `Device` réalise l'acquisition des données de contact et les convertit au format générique de l'intergiciel.

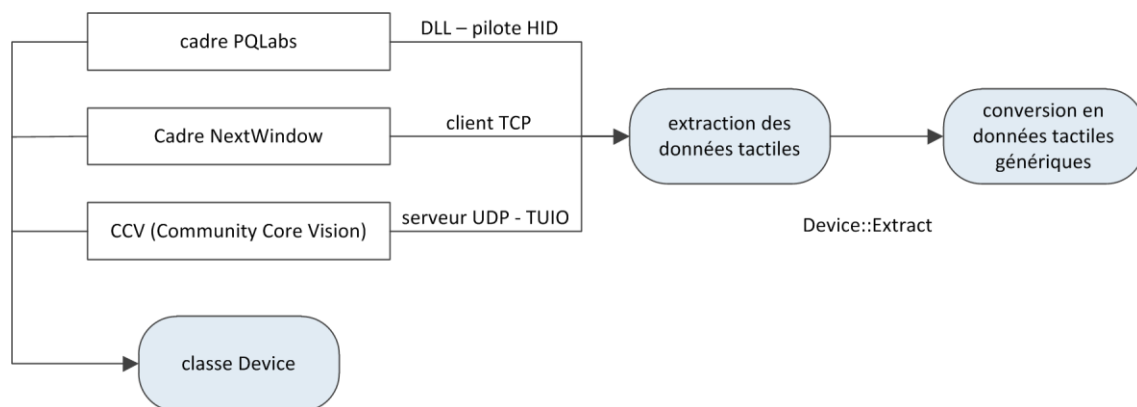


Figure 98 - Organigramme fonctionnel d'extraction et de conversion des données tactiles.

CCV : Protocole TUIO

Le protocole TUIO fournit à l'intergiciel les données tactiles par paquets OSC, sur le réseau local. Le protocole TUIO suit la structure suivante :

```
/tuo/[profil ] source application@adresse  
/tuo/[profil ] alive [liste des ID de session]  
/tuo/[profil ] set [liste des paramètres]  
/tuo/[profil ] fseq (ID image)
```

Le profil que nous utilisons pour la communication avec notre intergiciel est défini par la syntaxe suivante :

```
/tuio/2Dcur set s x y X Y m
```

s	:	ID de session temporaire	:	int32
x,y	:	position	:	int32
X,Y	:	vitesse	:	float32
m	:	accélération	:	float32

API PQLabs

Afin de récupérer les données tactiles issues du cadre optique *PQLabs*, l'intergiciel doit se connecter au serveur TCP par la méthode suivante :

```
int ConnectServer(const char *ip = "127.0.0.1")
```

Puis envoyer au serveur une requête afin de choisir le format de données attendu :

```
int SendRequest(const TouchClientRequest & request)
```

La structure `TouchClientRequest` est définie ci-dessous :

```
struct TouchClientRequest
{
    // le type de requête
    int type;

    // un identifiant unique pour l'application qui se connecte
    GUID app_id;

    // Nom du module de reconnaissance de geste
    char param[128];
};
```

Parmi tous les types proposés par l'API, nous avons choisi de récupérer les données brutes afin d'effectuer nos propres traitements et unifier ainsi les valeurs entre les différentes plateformes :

Type	value	Description
RQST_RAWDATA_ALL	0x0002	Le serveur envoie à l'application client l'intégralité des données brutes (<i>raw data</i>) lorsque l'utilisateur effectue un TouchDown, un TouchUp ou un TouchMove

Structure des données brutes :

```
struct TouchPoint
{
    Unsigned short    point_event;
    Unsigned short    id;
    int               x;
    int               y;
    Unsigned short    dx;
    Unsigned short    dy;
};
```

point_event : indique le type d'événement du point de contact concerné (TP_DOWN, TP_MOVE, TP_UP)

id : indique l'identifiant unique du point de contact concerné par l'événement

X : spécifie la position du point de contact sur l'axe des abscisses, en pixels

Y : spécifie la position du point de contact sur l'axe des ordonnées, en pixels

dx : spécifie la largeur de la zone de contact, en pixel

dy : spécifie la hauteur de la zone de contact, en pixel

API NextWindow

Pour récupérer les données tactiles, issues du cadre optique NextWindow, l'intergiciel doit se connecter au driver par la méthode suivante :

```
void DeviceConnectHandler(DWORD deviceID)
```

La fonction TouchPacketReceived() est appelée à chaque nouveau paquet de données :

```
void TouchPacketReceived (Dword deviceID,
                          Dword deviceStatus,
                          Dword packetID,
                          Dword touches,
                          Dword ghostTouches);
```

La méthode `GetTouch()` permet de récupérer les données tactiles selon la structure suivante :

```
struct typedef
{
    Dword        touchID;
    Dword        touchType;
    __int64      touchStart;
    Point_t      touchPos;
    Float        velocity;
    Float        acceleration;
    Float        touchArea;
    Dword        touchEventType;
    Dword        confidenceLevel;
    Dword        height;
    Dword        width;
}
```

<code>touchID</code>	:	indique l'identifiant unique du point de contact concerné par l'événement
<code>touchType</code>	:	indique le type de point de contact reporté (TT_TOUCH, TT_GHOSTTOUCH et TT_CENTROID)
<code>touchStart</code>	:	indique le moment exact de la création du point de contact
<code>touchPos</code>	:	indique les coordonnées x, y du point de contact
<code>touchArea</code>	:	indique la taille de la zone de contact
<code>touchEventType</code>	:	indique le type d'événement (TOUCH_DOWN, TOUCHING, TOUCH_UP)
<code>confidenceLevel</code>	:	indique le niveau de confiance de l'existence réelle du point de contact
<code>Height</code>	:	spécifie la largeur de la zone de contact, en pixel
<code>Width</code>	:	spécifie la hauteur de la zone de contact, en pixel

Intergiciel : la classe Device

La classe `Device` dont héritent les classes de périphériques contient les méthodes générales d'accès et de manipulation des données tactiles entrantes.

Instanciée en singleton, nous avons fourni à la classe `Device` les méthodes suivantes :

```
bool connect()
```

Crée une connexion client/serveur TCP (PQLabs), UDP (TUIO), ou permet de se connecter au périphérique HID (NextWindow). Renvoie `true` si la connexion réussit.

```
void clear()
```

Vide le cache des données tactiles placées en mémoire.


```
bool extract(std::list<TouchPoint>& pts)
```

Récupère les informations de points de contact listées par les périphériques, les convertit en données génériques.
Si de nouvelles données existent, la fonction renvoie `true`.

```
void lock() / void unlock()
```

Bloque ou débloquent l'accès au périphérique par l'usage d'un mutex afin d'éviter les problèmes de synchronisation des points de contact actifs.

```
static Device*buildDevice(Device::DEVICE type)
```

Basée sur la structure de la classe `Abstract Factory`, cette méthode permet d'instancier une classe pour le périphérique à utiliser. Renvoie une instance de cette classe castée par un pointeur de la classe `Device` : `InputPQLabs / InputNextWindow / InputWall`.

Intergiciel : format des données tactiles

La conversion des données tactiles dans un format générique défini par la classe `TouchPoint` permet à l'intergiciel d'avoir accès aux mêmes informations tactiles, indépendamment du périphérique connecté. Nous avons établi une liste de propriétés que nous avons jugées essentielles, suite au modèle d'interaction multi-tactile défini précédemment, pour le design d'interaction permis par ce type de dispositifs.

D'une part, la méthode `extract` de la classe `Device` permet de récupérer les propriétés communes à tous les dispositifs :

<code>int</code>	<code>_id</code>	Identifiant
<code>int</code>	<code>_timestamp</code>	Date de génération
<code>Point2D<float></code>	<code>_pos</code>	Position
<code>Point2D<float></code>	<code>_size</code>	Taille du blob
<code>TOUCH_TYPE</code>	<code>_type</code>	Type (DOWN, MOVE, UP)

Ensuite, à partir de ces données générales, les propriétés suivantes sont calculées :

<code>float</code>	<code>_X, _Y</code>	Vecteurs vitesse
<code>int</code>	<code>_m</code>	Accélération

Enfin, les propriétés suivantes sont dépendantes de l'environnement

<code>static Point2D<int> _screen</code>	Résolution d'écran
<code>char _idApp</code>	Identifiant de l'application associée

Si le périphérique ne fournit pas ces informations, nous avons défini un système de coordonnées sur un repère orthonormé qui s'inspire de celui intégré au protocole TUIO. Représentées par des valeurs normalisées, codées en virgule flottante (`float`) comprises entre 0 et 1, x_C et y_C correspondent à la position relative des points de contact en fonction de la résolution du capteur. x et y correspondent aux valeurs absolues du point de coordonnées x_C et y_C en fonction de la résolution d'écran. Les unités des vecteurs vitesse X et Y correspondent au déplacement d'un point sur la totalité d'un axe en une seconde. Voici les formules appliquées au calcul des paramètres d'un point de contact C , de vitesse V (en unités par seconde) et d'accélération A , détectés par un capteur de résolution *largeur* × *hauteur* :

$$x = x_C \times \text{largeur} \quad y = y_C \times \text{hauteur} \quad V = \frac{\sqrt{(\Delta x)^2 + (\Delta y)^2}}{\Delta t}$$

$$X = \frac{\Delta x}{\Delta t} \quad Y = \frac{\Delta y}{\Delta t} \quad A = \frac{\Delta V}{\Delta t}$$

2.3. Architecture réseau

2.3.1. Système de communication asynchrone

Afin d'alléger le code à écrire par les développeurs côté application, nous avons transféré toute la partie de traitement des informations tactiles et des algorithmes de reconnaissance du geste (manipulation et gestuelle) de l'application vers l'intergiciel. Cette démarche présente d'autres avantages comme la simplification de l'API à développer pour tout nouveau langage ou programme, et la possibilité de réaliser des procédures de mise à jour de l'intergiciel sans devoir recompiler les applications. Dans l'idée de permettre à l'intergiciel de connaître le container ou l'objet de l'interaction ainsi que le répertoire des comportements et des gestes qui lui sont associés, nous avons mis en place un système de communication asynchrone entre l'intergiciel et l'application multi-tactile.

Pour décrire le fonctionnement général de l'environnement de développement, j'utiliserai un cas d'utilisation simple et concret de manipulation d'objet graphique. A l'écran sont représentées une dizaine d'images, de tailles différentes, et disposées de manière aléatoire. L'utilisateur a la possibilité de modifier les propriétés de ces images (position, taille,

rotation), de les mettre en plein écran en cliquant dessus, et de les disperser par une gestuelle imitant l'action de secouer.

Le diagramme ci-dessous représente le déroulement de l'interaction des points de contact à travers quatre étapes de coordination entre l'intergiciel et l'application multi-tactile :

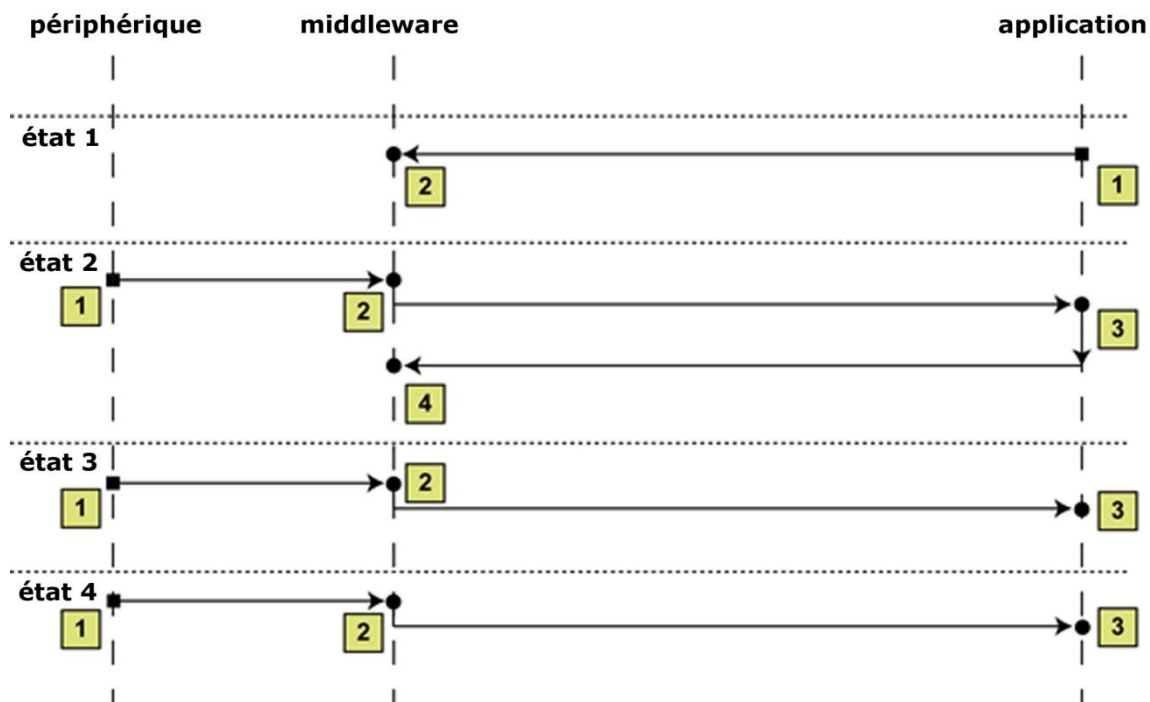


Figure 99 - Illustration de la synchronisation entre l'intergiciel et l'application

État 1 :

1. L'API client se connecte à l'intergiciel en TCP sur le port 3000 par défaut.
2. L'intergiciel enregistre l'application comme active grâce à l'identifiant envoyé par le client (THA).

État 2 :

1. L'utilisateur pose le doigt sur l'écran : le périphérique propage un événement de type `TouchDown` (nouveau point de contact) analysé par l'intergiciel.
2. L'intergiciel enregistre le point et le sauvegarde dans un tableau temporaire, en tant que point « orphelin », c'est-à-dire sans application, container ni objet encore associé, puis envoie à l'API client un message de type `TouchInfo`.

3. L'API client reçoit le paquet `TouchInfo` et en extrait les informations de position du point de contact. En fonction de ces coordonnées, l'API interroge la couche graphique (2D / 3D) et cherche l'objet et/ou le container situé sous ce point :
 - a. Si l'un ou l'autre possède un comportement multi-tactile ou gestuel, défini par le designer d'interaction, l'API construit un paquet de patrons d'interaction en renseignant les modalités et les paramètres de ces interactions.
 - b. Sinon, l'API crée un PPP vide.
4. Si b. est vérifié, l'API renvoie à l'intergiciel le PPP qui est analysé pour attribuer au point orphelin son lien de parenté avec l'objet graphique, le container, les PPP et l'application active.

État 3

1. L'utilisateur déplace le point de contact généré à l'état 2 et situé sur un objet graphique : le périphérique propage un événement de type `TouchMove` (point de contact déplacé) qui est récupéré par l'intergiciel.
2. L'intergiciel recherche si un point de contact déjà enregistré correspond à l'identifiant :
 - a. Si ce point n'existe pas encore (ce cas peut arriver lorsque le périphérique utilise le protocole UDP qui ne garantit ni la remise ni l'ordre des paquets délivrés). Dans ce cas, l'événement est mis en attente et son traitement est retardé jusqu'au prochain cycle, et ce jusqu'à la capture de l'événement `TouchDown` de ce point.
 - b. Si ce point existe : l'intergiciel connaît les interactions attendues par l'objet ou le container associé, il envoie alors la mise à jour des propriétés du point aux instances d'analyse et de traitement du geste (manipulation / gestuelle).
3. En fonction du PPP de l'objet ou du container, l'intergiciel envoie les informations d'interaction à l'application connectée, dans un message contextualisé événementiel (`InteractionEvent`) :
 - a. Si le PPP est de type « brut » (*raw data*), l'envoi s'effectue à chaque mise à jour des points de contact.

- b. Si le PPP concerne une interaction continue (généralement une manipulation), l'envoi s'effectue à chaque mise à jour définie par les algorithmes de transformation (agrandissement, déplacement, etc.)
 - c. Si le PPP concerne une interaction discrète (généralement une gestuelle), l'envoi s'effectue à la fin de la phrase gestuelle (état 4.2.), ou dès le moment où le module d'analyse du geste a identifié le geste en cours de réalisation.
4. L'API reçoit l'événement d'interaction et l'analyse afin d'identifier l'objet ou le container concerné. En les connaissant, l'API client propage un événement contenant le type d'événement (ex : déplacement) et ses propriétés (positions relatives, absolues, etc.)

État 4

1. L'utilisateur soulève le doigt de l'écran : le périphérique propage un événement de type `TouchUp` (point de contact supprimé) qui est récupéré par l'intergiciel.
2. Comme à l'état 3.2. l'intergiciel recherche si un point de contact déjà enregistré correspond à l'identifiant :
 - a. Si ce point n'existe pas (possible si perte du paquet `TouchDown` en UDP), l'intergiciel ne tiendra pas compte de l'événement et tous les événements mis en cache, relatifs à ce point de contact, sont supprimés.
 - b. Si ce point existe mais qu'il reste d'autres points associés au même objet ou au même container, l'intergiciel effectue, en fonction des PPP, une correction des coordonnées des points de référence.
 - c. Si ce point existe et qu'il ne reste plus aucun point associé au même objet ou au même container, l'intergiciel propage l'événement de fin de phrase gestuelle, supprime les occurrences d'analyse d'interaction, lance l'analyse des interactions discrètes éventuelles, et envoie ces informations sous la forme d'événements d'interaction `InteractionEvent` à l'API client.
3. L'API reçoit les événements d'interactions discrètes, les analyse et propage les événements des objets et/ou containers concernés à l'application. La fin de la phrase gestuelle est déclarée par la syntaxe balisée suivante :

```
<pState>[id_objet]:end</pState>
```

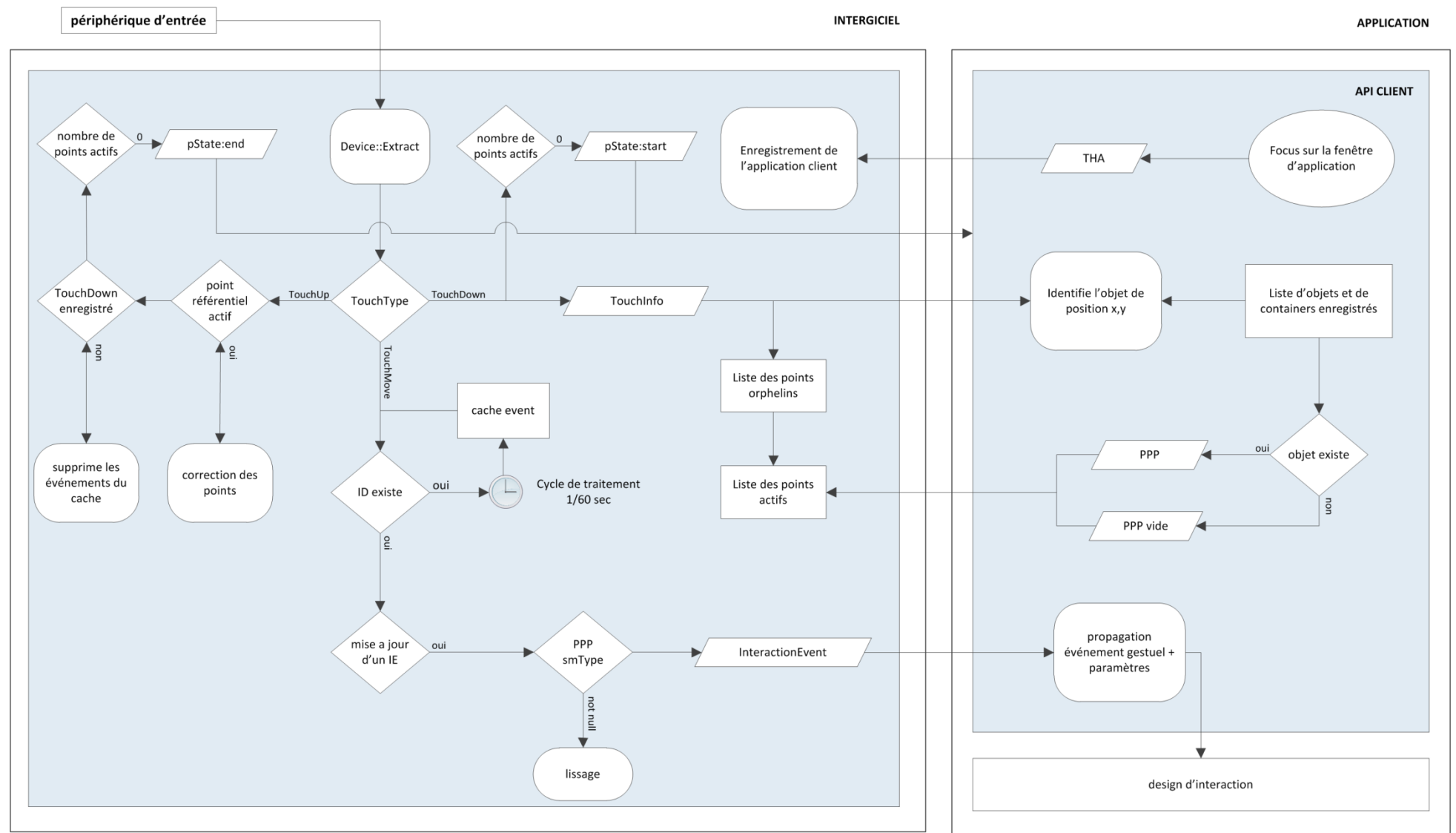


Figure 100 - Cycle de gestion des événements tactiles.

2.3.2. Paquet « TouchInfo »

Les paquets `TouchInfo` sont les premiers messages échangés entre l'intergiciel et l'API client, après l'envoi du `THA` d'enregistrement de l'application. Ces échanges utilisent le protocole TCP que nous avons privilégié pour la garantie de la remise et de l'ordre des paquets, conditions essentielles pour le bon fonctionnement de notre environnement de développement. La fréquence d'envoi des paquets n'étant pas supérieure à 1/60 de seconde, ce protocole de communication s'est avéré tout à fait approprié pour une installation locale (l'intergiciel étant installé sur le même système que l'application).

Les paquets `TouchInfo` sont envoyés par l'intergiciel lorsque le périphérique détecte un nouveau point de contact (cf. état 2.2). Les informations de ce point sont converties en données tactiles génériques, stockées dans un tableau de points orphelin et envoyées à l'application client pour obtenir des informations liées à l'objet et/ou au container concerné par l'interaction (PPP). Les données relatives à ce nouveau point de contact sont converties au format défini par la classe `TouchInfo`, format également utilisé pour la sérialisation des messages de type `rawdata` utilisés pour les `InteractionEvent` (cf. Chapitre VI - 2.3.4.). Envoyés à l'API client, les messages `TouchInfo` sont sérialisés selon la syntaxe balisée suivante :

```
<touchInfo>[TOUCH_TYPE]:[x]:[y]:[X]:[Y]:[A]:[dx]:[dy]</touchInfo>
```

`TOUCH_TYPE`

- D : correspond à un nouveau point de contact.
- M : correspond au déplacement d'un point de contact.
- U : correspond à la suppression d'un point de contact.
- x : indique la position relative du point de contact sur l'axe des abscisses (horizontal).
- y : indique la position relative du point de contact sur l'axe des ordonnées (vertical).
- X : vecteur vitesse sur le plan horizontal.
- Y : vecteur vitesse sur le plan vertical.
- A : accélération de la vitesse.
- dx : largeur de la zone de contact
- dy : hauteur de la zone de contact

2.3.3. Paquet de Patrons et de Paramètres d'interaction (PPP)

Les Paquets de Patrons et de Paramètres d'interaction multi-tactiles et gestuels sont les deuxièmes paquets (dans l'ordre chronologique) échangés par l'intergiciel et l'API client. Ils sont envoyés par l'API client lorsque que l'objet, le container et les PPP ont été identifiés par l'API (cf. état 2.2). Même si aucun PPP n'est associé à l'objet ou au container, le message est composé de deux parties :

- Les informations liées aux containers. Dans le cas d'application 2D, les containers concernent de simples objets graphiques ou la fenêtre générale de l'application. Dans le cas d'application 3D, les containers concernent les différentes vues caméra (*viewport*), ou la fenêtre générale de l'application.
- Les informations liées aux objets. Pour une application 2D, il s'agit de tout objet situé sur la couche graphique et déterminé par les coordonnées en deux dimensions du `TouchInfo`. Pour une application 3D, il s'agit de tout modèle situé dans l'univers 3D et détecté par lancer de rayon à partir des coordonnées en deux dimensions.

Les objets graphiques et les containers sont référencés par l'API dans deux listes distinctes qui sont interrogées pour obtenir leur identifiant. Envoyé à l'intergiciel, les messages PPP sont sérialisés selon la syntaxe balisée suivante :

```
<ppp>[ID du viewport]:[PPI]*n|[ID de l'objet]:[PPI]*n </ppp>
```

ID_VIEWPORT

Viewport unique ID.

ID_OBJECT

Object unique ID .

PPI*n

Liste des Patrons et des Paramètres d'Interaction attribués à l'objet ou au container.

2.3.4. Patron et Paramètres d'Interaction (PPI)

Les patrons (description) et les paramètres d'interaction (PPI) définis par le designer d'interaction sont encapsulés dans les paquets PPP. Ils décrivent l'ensemble des interactions

attendues par l'application et les paramètres de leur exécution. Afin de simplifier la tâche du développeur côté application, nous avons défini une liste d'interactions communément utilisées dans les interfaces multi-tactiles et gestuelles. Cet inventaire n'est pas exhaustif mais correspond à la palette d'interactions que nous avons sélectionnées pour la première version de l'intergiciel. Déclinées en interactions de manipulation ou en gestuelle, suite aux définitions apportées par notre modèle d'interaction, les PPI suivent la syntaxe balisée suivante :

`<PPI>[PATTERN_TYPE]:[PATTERN_DESC]</PPI>`

PATTERN_NAME	PATTERN_TYPE	PATTERN_DESC
TouchDown	DWN	<nMin:nMax>
TouchTap	TAP	<nMin:nMax:timeInterval:delta>
TouchDoubleTap	DTP	<nMin:nMax:timeInterval:delta>
TouchUp	TUP	<nMin:nMax>
TouchMove	MOV	<nMin:nMax:smType:smScope:delta>
TouchZoom	ZOM	<nMin:nMax:smType:smScope:delta>
TouchRotate	ROT	<nMin:nMax:smType:smScope:delta>
TouchRaw	RAW	<nMin:nMax>
GestureSelect	GST	<nMin:nMax:delta:minArea>
GestureWrite	GWT	<nMin:nMax:anFilterInc:anFilterExc>
GestureDraw	GDW	<nMin:nMax:formFilterInc:formFilterExc>

Patrons d'interaction (PATTERN_NAME)

Les patrons d'interaction (*interactive pattern*) sont une liste d'interactions tactiles et gestuelles, mises à disposition du designer d'interaction qui, de manière simple et rapide, peut créer des interactions plus riches sans être obligé d'en écrire les algorithmes les plus complexes. Ces patrons d'interaction sont encapsulés dans les PPI (liste des objets + paramètres) et envoyés de l'application à l'intergiciel par l'intermédiaire des PPP (paquets d'information envoyés lors d'un nouveau point de contact) pour activer les modules d'analyse des interactions correspondantes. Lorsque les conditions de la réalisation d'une interaction associée à un objet ou à container sont vérifiées, l'intergiciel renvoie à l'application un paquet `InteractionEvent` pour prévenir l'application de l'interaction en cours. Pour la première version de l'environnement de développement, nous avons réalisé une sélection de gestes de manipulation et de gestes disposant d'un fort degré de conventionnalité (ex: `GestureWrite`, `GestureDraw`, `GestureSelect`), ou appartenant désormais à des standards (`TouchZoom`, `TouchRotate`). Pour ne pas limiter

le designer d'interaction aux PPI existants, nous avons créé un patron d'interaction générique, qui contient l'intégralité des données de tout nouvel événement tactile, c'est le `TouchRaw`. De plus, l'API client ne fait que recevoir des événements tactiles simplifiés, accessibles par le développeur qui peut les utiliser à sa guise. Ainsi, un événement `TouchMove` par exemple, peut aisément être détourné pour régir une toute autre interaction que le déplacement d'un objet graphique.

`TouchDown`

L'intergiciel propage l'événement au début de la phrase gestuelle, lorsque l'utilisateur pose le doigt sur l'écran et qu'un nouveau point de contact est créé.

`TouchUp`

L'intergiciel propage l'événement lorsque l'utilisateur soulève son doigt de l'écran et qu'un point de contact existant est supprimé. Cet événement correspond à la fin de la phrase gestuelle lorsqu'il n'existe plus aucun autre point de contact associé à l'objet.

`TouchTap`

L'intergiciel propage l'événement lorsque l'utilisateur pose le doigt sur l'écran et ne le soulève qu'après une période définie par le paramètre `timeInterval`. Cette interaction est aussi connue sous le nom d'« appui long » ou « long Tap ».

`TouchDoubleTap`

L'intergiciel propage l'événement lorsque l'utilisateur pose le doigt deux fois sur l'écran de manière consécutive, en un temps inférieur à celui défini par le paramètre `timeInterval` et une différence de position inférieure à celle définie par le paramètre `delta`.

`TouchMove`

L'intergiciel propage l'événement lorsque l'utilisateur déplace son doigt à l'écran et modifie la position d'un point de contact existant. Il est possible de préciser le seuil minimum de déplacement nécessaire à l'activation de la propagation de l'événement grâce au paramètre `delta`.

`TouchZoom`

L'intergiciel propage l'événement lorsque l'utilisateur utilise deux doigts ou plus en augmentant ou en diminuant l'écart entre les points de contact. L'événement peut être utilisé pour modifier la taille d'une image ou le zoom d'une caméra 3D par exemple. Il est possible de préciser le seuil minimum d'écartement ou de pincement nécessaire à l'activation de la propagation de l'événement grâce au paramètre `delta`.

`TouchRotate`

L'intergiciel propage l'événement lorsque l'utilisateur utilise deux doigts ou plus en effectuant un mouvement de rotation des points de contact. L'événement contient des informations supplémentaires comme le centroïde de la rotation. Il est possible de préciser le seuil minimum de rotation nécessaire à l'activation de la propagation de l'événement grâce au paramètre `delta`.

TouchRaw

L'intergiciel propage l'événement lorsqu'un point de contact est créé, déplacé ou supprimé. Le format de l'événement est identique au format `TouchInfo`.

GestureSelect

L'intergiciel propage l'événement lorsque l'utilisateur effectue un mouvement (cercle, rectangle, etc.) de sélection, c'est-à-dire lorsque la position du point de départ est suffisamment proche du point d'arrivée (`delta`), et si l'aire de la zone de sélection est supérieure au paramètre `minArea`.

gestureWrite

L'intergiciel propage l'événement à la fin de la phrase gestuelle, lorsque l'utilisateur a soulevé tous ses doigts de l'écran et qu'il a procédé à la reconnaissance d'écriture. Sont inclus pour le moment les caractères alphabétiques et les chiffres. Il est possible de passer en paramètre une liste de caractères à exclure ou à inclure.

gestureDraw

L'intergiciel propage l'événement à la fin de la phrase gestuelle, lorsque l'utilisateur a soulevé tous ses doigts de l'écran et qu'il a procédé à la reconnaissance de forme. Les formes disponibles dans la version actuelle de l'intergiciel sont :

"RND" :	rond
"RCT" :	rectangle
"CRC" :	cercle
"CRX" :	croix
"SWL" :	balayage vers la gauche (swap left)
"SWR" :	balayage vers la droite (swap right)
"SWD" :	balayage vers le bas (swap down)
"SWT" :	balayage vers le haut (swap top)
"SCM" :	secouement
"TFD" :	pose des dix doigts de la main mimant la pose sur un clavier

Description des paramètres (`PATTERN_DESC`)

nMin

Entier indiquant le nombre minimum de points de contact pour activer l'interaction.

nMax

Entier indiquant le nombre maximum de points de contact pour activer l'interaction.

timeInterval

Entier indiquant un intervalle de temps en ms, utilisé notamment pour le `TouchTap` (intervalle de temps entre le `TouchDown` et le `TouchUp`).

smType

Chaîne de caractère indiquant le type de lissage de points effectué sur le patron d'interaction :

"none"
"mean:[puissance]"
"kalman" (prévu pour la prochaine version)

smScope

Entier indiquant la taille du tableau de points à utiliser pour le lissage.

delta

Nombre à virgule flottante indiquant un seuil minimum ou maximum pour l'activation du patron d'interaction.

anFilterInc

Liste de caractères alphanumériques, séparés par une virgule, à inclure dans la reconnaissance d'écriture. Par défaut, la liste contient tous les caractères alphabétiques et tous les chiffres. [0-9] correspond à l'ensemble des chiffres. [a-z] et [A-Z] aux lettres alphabétiques.

anFilterDec

Liste de caractères alphanumériques, séparés par une virgule, à exclure de la reconnaissance d'écriture. Par défaut, cette liste est vide. [0-9] correspond à l'ensemble des chiffres. [a-z] et [A-Z] aux lettres alphabétiques.

formFilterInc

Liste de formes, séparées par une virgule, à inclure dans la reconnaissance de formes. Par défaut, la liste contient toutes celles définies dans le dictionnaire de formes.

formFilterDec

Liste de formes, séparées par une virgule, à exclure de la reconnaissance de formes. Par défaut, cette liste est vide.

2.3.5. Paquet « InteractionEvent »

Les paquets `InteractionEvent` sont envoyés par l'intergiciel et analysés par l'API client pour signifier à l'application qu'un des événements tactiles ou gestuels, définis par le PPI, est en cours de réalisation (interaction continue) ou réalisé (interaction discrète). Les `InteractionEvent` contiennent la référence de l'objet concerné, le type d'interaction et divers paramètres liés à la manipulation ou à la gestuelle. Les `InteractionEvent` sont définis par la syntaxe balisée suivante :

```
<IE>[ID viewport]:<[PATTERN_TYPE] | [IE_PARAMS]>*n:[ID objet]: <[  
    PATTERN_TYPE] | [IE_PARAMS]>*n</IE>
```

ID viewport

Identifiant du container, initialement enregistré par l'intergiciel.

PATTERN_TYPE

Le type d'événement actif

<...>*n

Liste des événements actifs et des paramètres de chaque élément

Description des `InteractionEvent` :

InteractionEvent	PATTERN_TYPE	IE_PARAMS
TouchDown	DWN	idTouches:x:y:X:Y
TouchTap	TAP	idTouches:x:y:X:Y
TouchDoubleTap	DTP	idTouches:x:y:X:Y
TouchUp	TUP	idTouches:x:y:X:Y
TouchMove	MOV	idTouches:x:y:X:Y:V:x_rel:y_rel
TouchZoom	ZOM	idTouches:x:y:zoom:zoom_rel
TouchRotate	ROT	idTouches:x:y:rotation:rotation_rel
TouchRaw	RAW	Format TouchInfo
GestureSelect	GST	idTouches:x:y:box
GestureWrite	GWT	idTouches:x:y:recChar
GestureDraw	GDW	idTouches:x:y:recForm

IE_PARAMS

idTouches : indique la liste des identifiants des points impliqués dans l'interaction.

x : indique la position (entre 0 et 1) de l'interaction sur l'axe horizontal. Si l'interaction implique plusieurs points de contact, la position calculée est celle du barycentre de l'interaction.

y : indique la position (entre 0 et 1) de l'interaction sur l'axe vertical. Si l'interaction implique plusieurs points de contact, la position calculée est celle du barycentre de l'interaction.

X : vecteur vitesse sur le plan horizontal du barycentre de l'interaction.

Y : vecteur vitesse sur le plan vertical du barycentre de l'interaction.

V : indique la vitesse relative de déplacement du barycentre de l'interaction.

x_rel : indique la différence de position horizontale du barycentre de l'interaction par rapport à la dernière valeur enregistrée pour cet événement.

y_rel : indique la différence de position verticale du barycentre de l'interaction par rapport à la dernière valeur enregistrée pour cet événement.

`timeInterval` : indique un intervalle de temps. Cette information concerne principalement les événements `TouchTap` et `TouchDoubleTap`.

`zoom` : indique la distance actuelle des points de référence par rapport à la distance initiale (lors de l'activation du patron d'interaction).

`zoom_rel` : indique la différence de distance entre les points de référence par rapport aux dernières valeurs enregistrées pour cet événement.

`rotation` : indique l'angle de rotation trigonométrique actuel des points de référence, en degré, par rapport à l'axe vertical et dans le sens anti-trigonométrique (*dextrorsum*).

`rotation_rel` : indique la différence de rotation des points de référence, en degré, par rapport aux dernière valeurs enregistrées pour cet événement.

`bBox` : indique la position et la taille du rectangle correspondant à la boîte de délimitation en deux dimensions issue de la sélection.

`recChar` : indique le caractère au format Unicode reconnu par le module de reconnaissance de caractère.

`recForm` : indique la forme reconnue par le module de reconnaissance parmi celles définies dans la bibliothèque de l'intergiciel.

2.4. Traitement et lissage

2.4.1. Correction des points et delta

Au cours de ce chapitre, nous avons mis en évidence la nécessité de considérer la phrase gestuelle dans son ensemble, de la phase de préparation à la phase de rétractation et tout particulièrement l'évolution de sa phase nucléaire. L'analyse des enregistrements effectués sur le mur multi-tactile nous a donné de précieux renseignements sur le cours kinésique du geste et sur l'évolution des points de contact de référence utilisés pour l'analyse des interactions. Les critères de sélection des points de références sont multiples. Cependant, deux cas de figure sont communs à la majorité des interactions :

- Les points de référence sont déterminés en fonction de leur ancienneté (`TouchMove`) : le point de référence correspond au point le plus ancien.
- Les points de référence sont déterminés en fonction de leur position (manipulations à 2 points et plus) : les points de référence correspondent aux points les plus distants.
- Les points de référence correspondent à l'ensemble des points assignés à un même objet (principe en cours d'évaluation).

Dans un premier temps, nous avons fait l'observation que les points de référence, en passant d'un doigt à un autre au cours de la phrase gestuelle, entraînaient des fluctuations importantes si, comme c'est le cas des API multi-tactiles existantes, le système ne corrige pas les écarts de valeur lors de la permutation de ces points. En effet, imaginons que le report d'un événement de déplacement `TouchMove` renvoie par défaut les coordonnées et la position relative (à la précédente) du point de contact p_0 . Si à un moment de la phrase gestuelle, ce point est supprimé et remplacé par le point de contact p_1 (point le plus ancien après p_0), la modification du point de référence entraînera un changement soudain de la position reportée. Malgré la présence de lissage de points, cette variation est particulièrement visible dans le cas d'interactions de manipulation (déplacement, agrandissement, rotation d'objets graphiques).

Pour remédier à ce phénomène, nous avons intégré à l'intergiciel un système de correction des coordonnées reportées par les points de référence. Pour l'ensemble des patrons d'interaction de manipulation, impliquant de 1 à n doigts, nous ajoutons aux coordonnées xy de tout nouveau point de référence de substitution, la différence de position entre ce nouveau point et l'ancien, et ceci pour l'ensemble des substitutions présentes au cours d'une même phrase gestuelle (tant qu'il existe des points de contact associés à un même objet ou container).

Soit une suite n , de points de références successifs $P_{i(i=1...n)}$ assignés à la même interaction. Soit $P'_{i(i=1...n)}$ les points de référence au moment de leur dernière position avant chaque substitution. Soit R le repère d'origine M et n le nombre de substitutions. Les nouvelles coordonnées du point P à l'instant t peuvent être décrites par la formule vectorielle suivante :

$$\overrightarrow{MP_{n(t)}}_{corrigé} = \overrightarrow{MP_{n(t)}} - \sum_{i=1}^{n-1} (\overrightarrow{P'_i P_{i+1}})$$

Simplifiée par la relation de Chasles :

$$\overrightarrow{MP_{n(t)}}_{corrigé} = \overrightarrow{OP_{n(t)}} - \overrightarrow{P'_1 P_n}$$

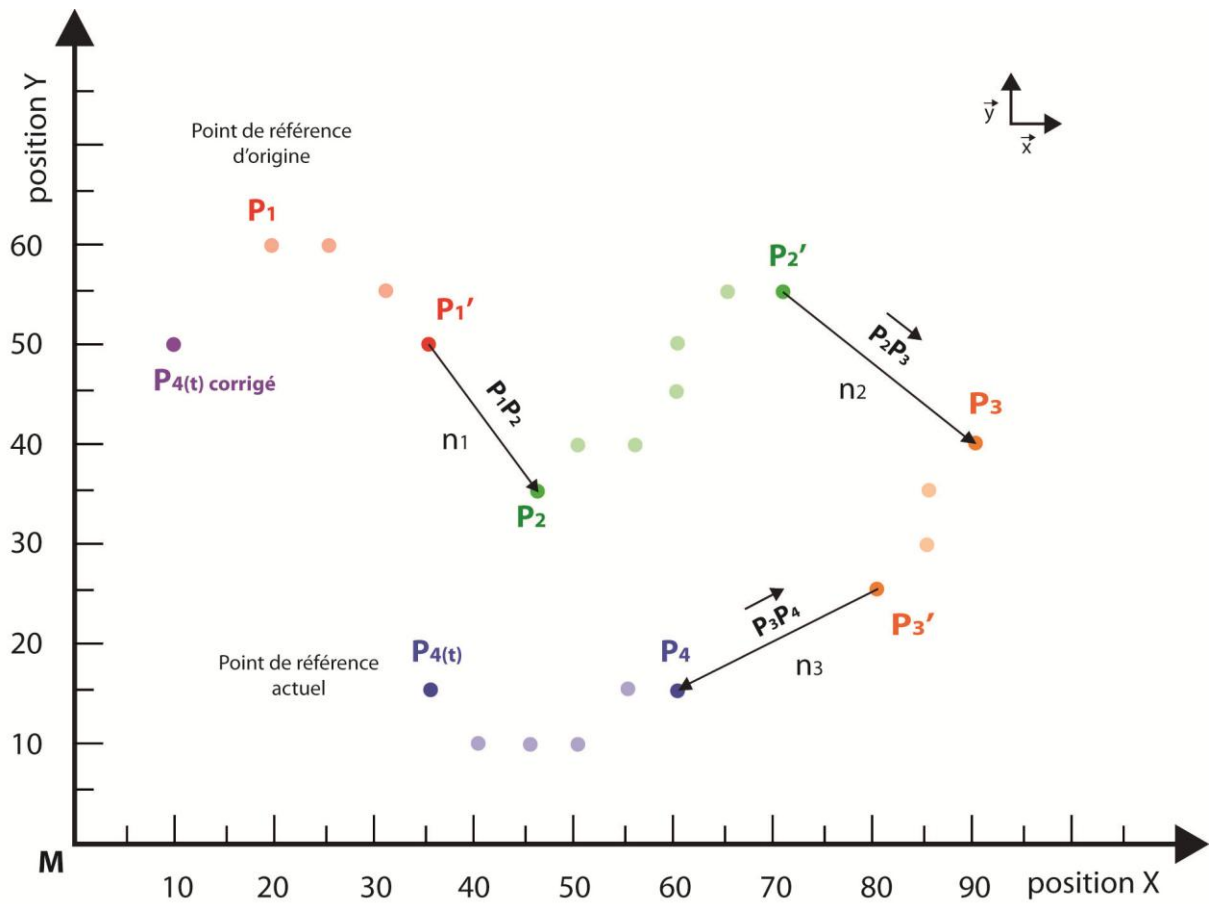


Figure 101 - Illustration simplifiée de l'estimation des coordonnées corrigées d'un point de référence P à l'instant t, suite à n substitutions du point de référence.

Dans l'exemple ci-dessus, il est possible de calculer la position corrigée du point P_4 de coordonnées $x=35$ et $y=15$ à l'instant t, suite aux trois changements de points de référence X, de la manière suivante :

$$\overrightarrow{P_M P_{4(t)} \text{ corrigé}} = \overrightarrow{P_M P_{4(t)}} - (\overrightarrow{P_1' P_2} + \overrightarrow{P_2' P_3} + \overrightarrow{P_3' P_4}) = \overrightarrow{P_M P_{4(t)}} - \overrightarrow{P_1' P_4}$$

$$\overrightarrow{P_M P_{4(t)} \text{ corrigé}} = (35\vec{x} + 15\vec{y}) - (25\vec{x} - 35\vec{y}) = 10\vec{x} + 50\vec{y}$$

Cette correction, en amont, entraîne la correction de toutes les valeurs calculées à partir de ces points de référence, comme le barycentre B des interactions suivantes :

TouchDown {nMin : 2}

TouchUp {nMin : 2}

TouchTap {nMin : 2}

TouchDoubleTap {nMin : 2}

TouchMove {nMin : 2}

TouchZoom {nMin : 2}

TouchRotate {nMin : 2}

TouchMove {nMin : 2}

Dont les coordonnées sont calculées selon la fonction vectorielle de Leibniz suivante :

$$\vec{f}(M) = \sum_{i=1}^n p_i \overrightarrow{MP_i}$$

où $P_{i(i=1...n)}$ est une famille de n points et $p_{i(i=1...n)}$ une famille de n scalaires.

2.4.2. Lissage de points

Le lissage des points de contact est une fonctionnalité que nous avons intégrée à l'intergiciel suite aux problèmes d'imprécision et de micro mouvements de l'utilisateur relevés lors des phases d'observation et d'analyse. Afin d'accorder au designer d'interaction une plus grande liberté, nous lui donnons la possibilité de déterminer les objets concernés et le type de lissage appliqué à ces objets.

La définition de ces propriétés se fait lors de l'enregistrement des patrons d'interaction en précisant le paramètre `smType` du PPP. Le lissage s'effectue sur un tableau de points dont la longueur dépend du paramètre `smScope`, défini également lors de la création du PPP.

Lissage par moyenne mobile pondérée ("mean:[puissance]")

La moyenne pondérée, ou glissante, est un type de moyenne que nous avons utilisé pour analyser la série temporelle des coordonnées ou des propriétés tactiles que l'intergiciel renvoie à l'application afin de lisser ces valeurs ou de limiter l'influence des points aberrants en réduisant les fluctuations transitoires.

La moyenne mobile pondérée est une moyenne qui permet d'utiliser des coefficients et de donner des poids distincts à chaque valeur du calcul. Les poids attribués à chaque terme décroissent linéairement, le plus récent possédant un poids n et le plus ancien ayant un poids unitaire.

La définition mathématique de cette moyenne mobile pondérée M , pour n valeurs de x à lisser et de pondération p , s'écrit de la manière suivante :

$$\bar{x}_M = \frac{np_M + (n-1)p_{M-1} + \dots + 2p_{(M-n+2)} + p_{M-n+1}}{n + (n-1) + \dots + 2 + 1}$$

Le dénominateur N correspond à la somme des valeurs de n , c'est donc un nombre triangulaire qui peut également être calculé grâce à la formule suivante :

$$N = \frac{n(n+1)}{2}$$

Toutefois, en appliquant ce filtre, avec un degré de décroissance linéaire des poids appliqué à chaque échantillon de valeur, notamment aux coordonnées x, y de la position d'un objet déplacé en fonction des coordonnées du barycentre d'une interaction de type `TouchMove`, nous avons remarqué, malgré la qualité du lissage réalisé, un effet de latence généré par ce filtre. Il nous est donc apparu nécessaire de l'appliquer avec un degré de décroissance exponentiel des poids appliqués à chaque échantillon afin d'accorder une influence plus importante aux dernières valeurs ajoutées à la liste de points et atténuer plus rapidement l'impact des valeurs plus anciennes. En ajoutant un coefficient de puissance α tel que $p_M = M^\alpha$. Avec $\alpha = 2$ et $n = 15$ périodes où une période correspond à 1/60 seconde, nous avons trouvé un juste équilibre entre l'efficacité du lissage et le temps de latence généré par l'algorithme.

3. Implémentation des API

Dans ce chapitre, nous nous intéresserons essentiellement aux méthodes de développement que nous avons simplifiées pour le designer d'interaction ou l'intégrateur. En effet, nous souhaitons proposer des interfaces de programmation accessibles par un développeur non initié aux contraintes et aux logiques propres au développement d'applications multi-tactiles. L'architecture logicielle et les diagrammes de classes des API restent à la discrétion de l'entreprise Virtual-IT, qui possède les droits de propriété industrielle.

3.1. La classe SorClient

La classe SorClient correspond à la classe principale de l'API client. Restreinte à une seule instantiation (singleton), elle permet d'initialiser le client TCP et les différents paramètres de l'application multi-tactile. Elle est commune aux interfaces de programmation (API) AS3 pour Adobe Flash et C# pour Unity 3D.

Elle dispose des méthodes publiques suivantes :

`init(scope)`

Initialise l'API client et lance la procédure de connexion avec l'intergiciel. N'existe qu'en AS3. La variable scope permet de transmettre la référence de la scène (Stage).

`Connect(hote, port)`

Tente de se connecter à l'intergiciel selon l'adresse et le port donnés en paramètres. Renvoie -1 si la connexion échoue.

Elle dispose des propriétés publiques suivantes :

`name`

Spécifie le nom de l'application client envoyé à l'intergiciel à travers le THA (*Tactile Header Application*). Si non renseignée, une valeur aléatoire est générée.

`host`

Spécifie l'adresse IP ou le nom d'hôte à utiliser pour la connexion avec l'intergiciel. La valeur par défaut est 127.0.0.1.

`port`

Spécifie l'adresse IP ou le nom d'hôte à utiliser pour la connexion avec l'intergiciel. La valeur par défaut est 5151.

`useShapeFlag(AS3 uniquement)`

Indique si le test de collision doit porter sur les pixels réels de l'objet (true) ou sur le cadre de délimitation (false).

`autoReconnect`

Spécifie si le client doit se reconnecter si la connexion est perdue.

`reconnectDelay`

Spécifie l'intervalle minimum entre deux tentatives de connexion.

`mouseEmulation`

Spécifie si le client doit indiquer à l'intergiciel d'émuler la souris.

mouseHide

Spécifie si le client doit masquer le curseur de la souris.

mouseEventReport (AS3 uniquement)

Spécifie si le client propage des événements de type `MouseEvent` pour rendre compatible des applications traditionnelles.

debugMode

Affiche sur la vue utilisateur les informations de connexion et la représentation des points de contact.

Sous Adobe Flash - AS3

```
private function initSorClient():void
{
    SorClient.name = "Application X v1.0.0";
    SorClient.host = "192.168.0.101";
    SorClient.port = 5152;
    SorClient.useShapeFlag = true;
    SorClient.autoReconnect = false;
    SorClient.reconnectDelay = 2000 ;
    SorClient.mouseEmulation = true;
    SorClient.mouseHide = true ;
    SorClient.mouseEventReport = true;
    SorClient.debugMode = true;
    SorClient.init(this);
}
```

Sous Unity 3D – C#

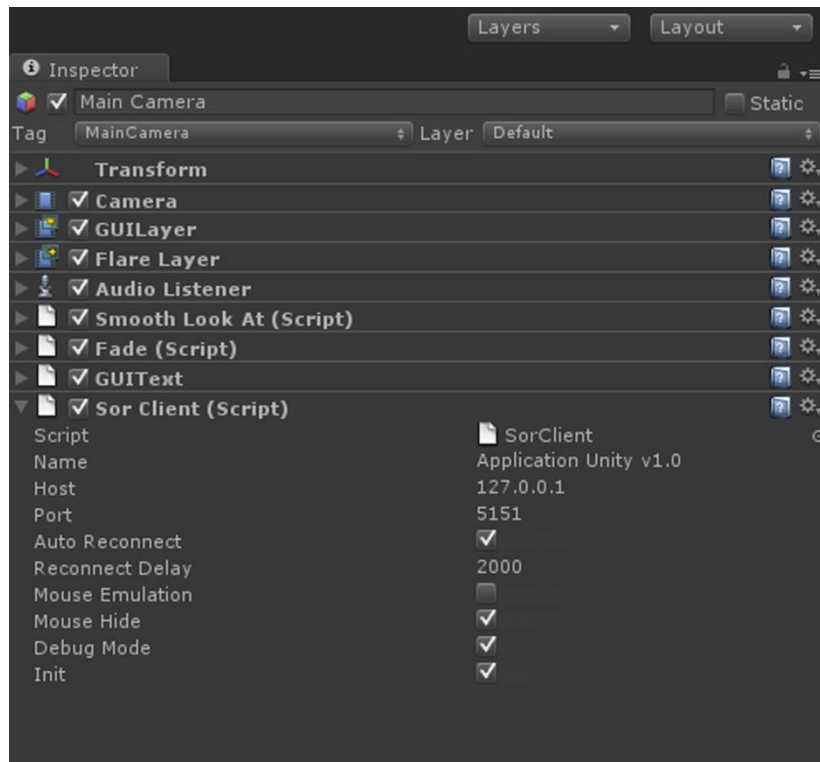


Figure 102 - Panneau de modification des paramètres du SorClient sous *Unity* 3D.

3.2. Interrogation de la couche graphique

Sous Adobe Flash - AS3

L'API est compatible avec l'ensemble des environnements de développements *Adobe* : *Flash*, *Flex* et *Air*. Nous avons utilisé cette technologie essentiellement pour la réalisation de prototypes logiciels en profitant de la simplicité d'interfaçage et d'animation permise par le logiciel *Adobe Flash*. Le langage *ActionScript*, étant interprété par une machine virtuelle (AVM : *ActionScript Virtual Machine*), nous avons conscience des performances limitées de cette plateforme. Toutefois, il nous a semblé intéressant de vérifier si en soulageant l'application client de la charge algorithmique nécessaire au traitement et à l'analyse des points de contact, nous optimisions ainsi les performances de l'application.

La classe `DisplayObjectContainer` de l'environnement de développement Flash AS3 est la classe de base de tous les objets graphiques susceptibles de servir de conteneur d'objets dans la liste d'affichage. Celle-ci gère tous les objets affichés dans le lecteur *Flash*. La classe `DisplayObjectContainer` permet d'organiser les objets d'affichage dans la liste d'affichage. Chaque objet `DisplayObjectContainer` possède sa propre liste d'enfants, qui organise l'ordre Z des objets. L'ordre z correspond à l'ordre d'empilement du premier plan vers l'arrière-plan.

L'enregistrement de tout nouvel objet graphique ou container multi-tactile se fait lors de l'ajout de cet objet sur la couche graphique. Nous avons créé la classe `MovieClipMt`, qui étend la classe `MovieClip`, en implémentant une méthode pour enregistrer l'objet dans la classe statique `MtObjectRef`. Cette classe offre des méthodes pour enregistrer ou supprimer ces objets, pour créer ou renvoyer leurs patrons d'interaction.

```
// Fonction appelée lors de l'ajout de l'objet graphique sur la couche graphique
private function onAddedToStage():void
{
    MtObjectRef.addObject(this);
}
```

Pour récupérer l'objet graphique ou le container (un objet graphique parent ou la scène) situé sous le point de contact, nous utilisons la méthode `hitTestObject()` de la classe `DisplayObject` qui permet de faire un test de collision entre un objet graphique et un point de coordonnées (x, y).

```
public static function onTouchDownInfo (pTouchPoint:Point, pShapeFlag:Boolean):void
{
    // Déclaration des variables
    var objectIdUnderPoint:int = -1;
    var highestDepth:int = -1;

    // Teste la collision avec l'ensemble des objets graphiques MT enregistrés
    for (var i:int=0; i<mtObjectArray.length; i++)
    {
        var curMc:MovieClipMt = mtObjectArray[i] as MovieClipMt;

        // Effectue un test de collision avec le point touchPoint, avec le paramètre
        // de cage de collision pShapeFlag
        if (curMc.hitTestPoint(pTouchPoint.x, pTouchPoint.y, pShapeFlag))
        {
            // Détermine l'objet situé au premier plan sous le point de contact
            if (curMc.parent.getChildIndex(curMc) > highestDepth)
            {

```

```

        objectIdUnderPoint = i;
        highestDepth = curMc.parent.getChildIndex(curMc);
    }
}

// Appelle la méthode buildPPP de la classe patternRefObject pour créer le PPP
var ppp:String = MtObjetRef.buildPPP(objectIdUnderPoint, _scope);

// Envoi le PPP à l'intergiciel
sendPPP(ppp) ;

}

```

Sous Unity - C#

Unity 3D est un environnement de développement qui permet d'exporter des scènes en 3D temps réel sous différentes plateformes. L'entreprise Virtual-IT s'est tournée vers ce logiciel pour sa capacité à publier pour ordinateur (*PC, Mac*), pour mobile (*iOs, Android*) et sur Internet.

La particularité de *Unity* 3D est d'utiliser des préfabriqués (ou *prefabs*) qui sont des composants facilitant le travail de l'intégrateur 3D. Il dispose de tout un ensemble de modèles et de comportements qu'il peut instancier, parfois sans même écrire une seule ligne de code. Les préfabriqués sont régis par des classes en C#, dont les propriétés publiques apparaissent distinctement dans le panneau de propriété du préfabriqué sur l'objet qui le reçoit. Pour rester dans la logique de Unity 3D, nous avons créé des classes d'événements multi-tactiles, qui une fois appliquées par glisser/déposer sur un objet, permet de sélectionner le ou les patrons d'interaction actifs et modifier leurs paramètres directement dans le panneau de propriété de l'objet.

Les objets 3D dans *Unity* sont classés selon des calques d'objets (*layers*), notamment utilisés par les caméras pour ne réaliser le rendu que de certaines parties de la scène. Nous avons profité de ce système de calque pour indiquer au moteur 3D les objets à inclure dans la liste des objets tactiles de l'application, en donnant à la propriété `layer` la valeur 31, la plus haute, généralement non assignée.

```
// Lors de l'initialisation de l'objet dans le moteur 3D

public void Awake()
{
    // Spécifie la valeur du calque attribué à l'objet
    gameObject.layer = 31;
}
```

Pour récupérer le container (vue caméra) et l'objet situé sous le point de contact, nous utilisons la méthode `Raycast()` de la classe `Physics` qui permet de faire un lancer de rayon et de récupérer les informations de collision avec les objets 3D.

```
void onTouchDownInfo(ref TouchPoint touchPoint)
{
    // Déclaration des variables
    Vector3 touchVector = new Vector3(0,0,0);
    Camera touchedCam;
    RaycastHit hitInfo = new RaycastHit();

    // Conversion du plan de coordonnées x, y
    touchPoint.y = 1f - touchPoint.y;

    // Création du vecteur 3D en fonction du point de contact touchPoint
    touchVector.x = (float) touchPoint.x * Screen.width;
    touchVector.y = (float) touchPoint.y * Screen.height;

    // Identification de la vue caméra active sous le point de contact (viewPort)
    if ( ( touchedCam = SceneObjectRequestor.getViewportUnderTP(tp, cameras)) == null)
        return ;

    // Fait une conversion du point de contact 2D en vecteur 3D
    ray = ccam.ScreenPointToRay(touchVector);

    // Fait un lancer de rayon et renvoie les résultats dans la variable hitInfo
    Physics.Raycast(ray, out hitInfo, Mathf.Infinity, 1 << 8);

    // Appelle la méthode buildPPP de la classe patternRefObject pour créer le PPP
    string ppp = MtObjetRef.buildPPP(hitInfo, touchedCam);

    // Envoi le PPP à l'intergiciel
    sendPPP(ppp) ;
}
```

3.3. Enregistrement des OCGM

Parmi les principaux objectifs de l'environnement de développement, nous souhaitons simplifier la méthode de création de comportements multi-tactiles et gestuels pour le développeur ou l'intégrateur, côté applicatif, qui n'a pas obligatoirement l'expérience de ce type de développement. D'une part, en déportant l'ensemble du traitement et de l'analyse des points de contact, nous avons simplifié et unifié les développements. D'autre part, nous avons créé des API en simplifiant la démarche d'enregistrement des patrons d'interaction.

Sous Adobe Flash - AS3

En nous inspirant de la logique de propagation d'événements existant pour les interactions de type `MouseEvent`, nous avons créé la classe `TouchEvent` et `GestureEvent`, qui étendent la classe `Event` et permettent d'enregistrer et de propager les événements correspondants. Ci-dessous, un exemple d'enregistrement de patrons d'interaction sur une image et sur la scène.

```
// Création d'un nouvel objet multi-tactile. type = "container" ou "object"
var image:MovieClipMt = new MovieClipMt(type);

// Ajout d'un objet Bitmap à l'image
image.addChild(bitmap);

// Ajout de l'objet graphique sur la liste d'affichage
addChild(image);

// ajout d'un patron d'interaction tactile TouchMove + TouchZoom + TouchRotate
image.addEventListener(TouchEvent.TOUCH_MOVE, onTouchMove, {nMin:1, nMax:10});

image.addEventListener(TouchEvent.TOUCH_ZOOM, onZoom, {nMin:2, nMax:10});

image.addEventListener(TouchEvent.TOUCH_ROTATE, onTouchRotate, {nMin:2,
nMax:10, smType:"mean:2", smScope:15});

// Ajout d'un patron d'interaction gestuel de reconnaissance de caractère
stage.addEventListener(GestureEvent.WRITE, onWrite, {nMin:1, nMax:1});
```

Ensuite, il suffit au designer d'interactions de créer la fonction appelée lors de l'événement et de définir les règles d'interactivité des objets graphiques en quelques lignes de code.

Exemple :

```
private function onZoom(pEvt:TouchEvent):void
{
    définition de pEvt : { idTouches, nbTouch, x, y, zoom, zoom_rel}

    // Pointage de l'objet à l'origine de l'événement
    var curImage:MovieClipMt = pEvt.currentTarget as MovieClipMt;

    // Agrandissement de l'image en fonction des données envoyées par l'intergiciel
    curImage.scaleX += pEvt.zoom_rel;
    curImage.scaleY += pEvt.zoom_rel;
}
```

Sous Unity 3D - C#

En nous inspirant de la logique des préfabriqués de l'environnement de développement propre à Unity 3D, nous donnons au développeur, ou à l'intégrateur, la possibilité de créer des patrons d'interaction très simplement, par l'intermédiaire de panneaux de propriétés des patterns glissés/déposés sur les objets 3D.

Il suffit ensuite de pointer une fonction par la propriété `callFunction`, un script par la propriété `component`, et un objet par la propriété `GameObjectTransform` pour propager l'événement tactile.

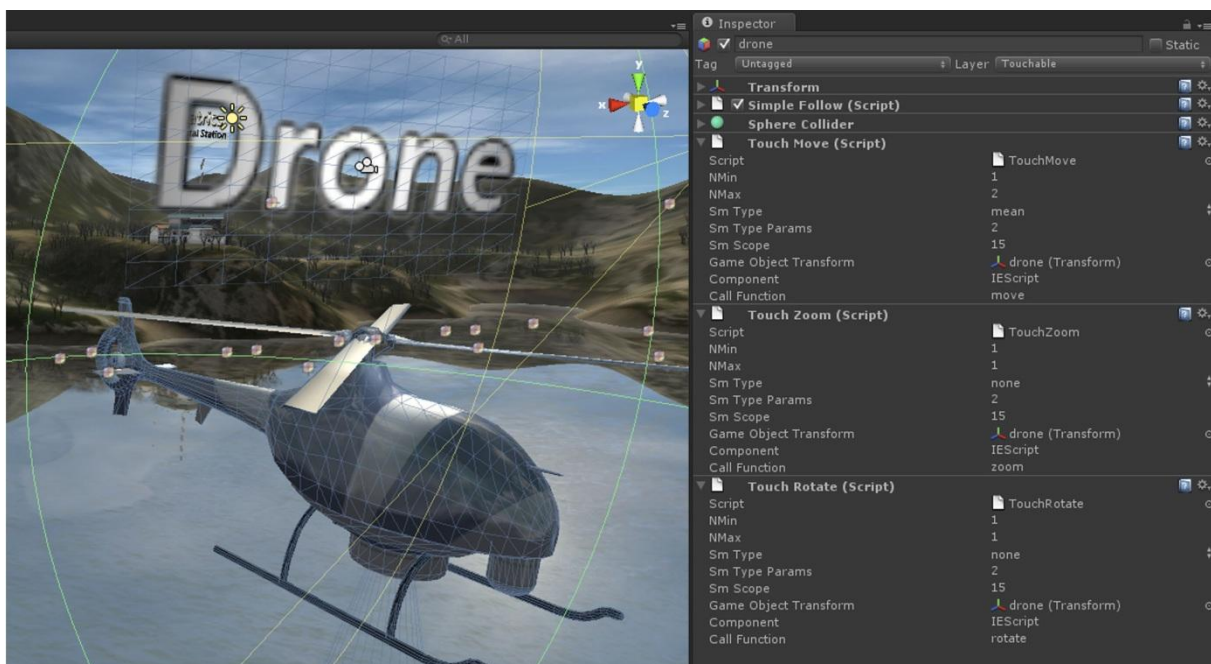


Figure 103 - Enregistrement des patrons d'interaction après glisser/déposer depuis la fenêtre de script, en modifiant les propriétés des patterns ainsi que les scripts et les fonctions à appeler lors de l'événement.

CONCLUSION

« Les choses les plus belles sont celles que souffle la folie et qu'écrit la raison. Il faut demeurer entre les deux, tout près de la folie quand on rêve, tout près de la raison quand on écrit. »

André Gide

Approche théorique

En m'engageant dans ce sujet de recherche, j'étais conscient de la richesse et de la diversité des contributions scientifiques existantes et de l'intérêt que j'avais d'approcher mon objet d'étude par une lecture pluridisciplinaire pour appréhender toute la portée des concepts qui lui étaient rattachés. Comme le souligne Edgar Morin [*La Tête bien faite*, 1999], la délimitation d'un champ disciplinaire, malgré son appartenance à un ensemble scientifique plus vaste, que Jean Piaget aimait appeler le « cercle des sciences », amène une richesse théorique indiscutable par la spécialisation de son langage, de ses techniques et de ses modes de production de connaissance. Toutefois, le cloisonnement disciplinaire, comporte également des risques que Morin formule en précisant que le terme désignait autrefois un petit fouet utilisé pour s'auto-flageller et que la discipline, en s'hyperspécialisant, devient un moyen de « flageller » celui qui s'aventure au-delà de son domaine de recherche. Il nous rappelle ainsi les limites théoriques de l'approche disciplinaire :

« On peut néanmoins dire très rapidement que l'histoire des sciences n'est pas seulement celle de la constitution et de la prolifération des disciplines, mais en même temps celle de ruptures des frontières disciplinaires, d'empiètements d'un problème d'une discipline sur une autre, de circulation de concepts, de formation de disciplines hybrides qui vont finir par s'autonomiser.²⁰² »

Pour faire évoluer ma question de départ et conduire mon analyse à identifier les problématiques soulevées par mon sujet de recherche, je n'ai pas hésité à faire quelques détours théoriques en empruntant les chemins de la psychologie, de la linguistique et de l'anthropologie. En faisant circuler les concepts véhiculés par les technologies multi-tactiles à travers cette grille de lecture, j'ai pu obtenir un cadre de compréhension général de mon objet d'étude. Ce n'est qu'à partir de ce moment qu'il m'a été possible de tracer les grandes lignes de cette analyse et de formuler mes hypothèses de départ. En déclarant que « *la science informatique n'est pas plus la science des ordinateurs que l'astronomie n'est celle des télescopes* », Edsger Dijkstra n'aurait pu décrire de meilleure manière ce champ disciplinaire pour évoquer la diversité et l'étendue des savoirs qu'il convoque. Au-delà des questions

²⁰² MORIN Edgar. *La Tête bien faite : Penser la réforme, reformer la pensée*. Seuil, 1999, 153 p. (Coll. L'histoire immédiate)

philosophiques, humanistes et sociologiques qui ont accompagné cette étude, nous avons pu constater combien les travaux sur la psychologie du développement et l'acquisition du savoir menés par les psychologues Jean Piaget et Jerome Bruner ont contribué aux fondements de la recherche en interaction homme-machine. Nous avons également mis en évidence l'intérêt de confronter les théories de la perception à notre objet d'étude en approchant le design d'interaction par les principes d'affordance, de visibilité et de feedback. Enfin, ce sont par les sciences du vivant et principalement grâce à la caractérisation des composantes cognitives, sensori-motrices et physiologiques du geste qu'il m'a été possible de définir de manière précise un modèle d'interaction théorique.

Dans un premier temps, en plaçant l'approche des interfaces multi-tactiles sous l'angle de l'instrument d'interaction, j'ai compris en quoi ces dispositifs se distinguent des interfaces haptiques et tangibles en s'inscrivant en différents points du continuum de virtualité. D'un côté du continuum, les instruments haptiques s'incarnent intégralement dans le réel de manière prothétique et procurent à l'utilisateur la conviction de la matérialité des objets virtuels par le réalisme des retours tactilo-kinesthésiques. Plus loin, les interacteurs mixtes, artefacts physiques de couplage entre l'espace réel et virtuel, profitent de l'affordance de leur forme physique et de leur spécialisation pour garantir le réalisme des interactions. Enfin, les interfaces tactiles déplacent la partie physique de l'instrument d'interaction vers la main de l'utilisateur, supprimant ainsi toute possibilité de préhension des instruments d'interaction ou des objets de la manipulation. En effet, les interfaces multi-tactiles étant caractérisées par l'absence d'instrument physique, la question de la perception et du feedback tactilo-kinesthésique s'est imposée très tôt dans l'orientation de ma recherche. Les conséquences de la disparition du périphérique d'entrée utilisateur, signifiant sa dématérialisation et sa dé-spécialisation, m'ont alors convaincu de la nécessité d'établir un modèle pour guider le design d'interaction multi-tactile.

En prolongeant la classification de Maria Karam et M.C. Schraefel, j'ai procédé à l'analyse physiologique de ces interfaces. D'une part, en examinant les principales technologies de perception tactile, j'ai mis en évidence l'influence des configurations techniques dans la qualité et la granularité des données tactiles capturées. D'autre part, cette analyse m'a fourni les éléments théoriques nécessaires à la définition des besoins et des objectifs concernant la fabrication des prototypes multi-tactiles. Enfin, l'étude sensorielle et motrice de la main et la définition d'une taxinomie générale du geste m'ont permis de relever les

limites du caractère « naturel » des interactions multi-tactiles. La manipulation des objets d'intérêt par des interactions directes, multiplexées dans l'espace, suggérait des modalités d'interaction plus « naturelles », en ce sens où elle engageait, au-delà de comportements innés, des formes de manipulation inspirées de l'expérience du monde physique et des formes de gestuelles issues de la conversation humaine non-verbale. Toutefois, en confrontant la richesse du canal gestuel à la morphologie spatiale et temporelle du geste en situation d'interaction multi-tactile, les limites conceptuelles du paradigme d'interface utilisateur « naturelle » ont fini par émerger.

Il semble en effet nécessaire de reconsidérer l'implication des formes « naturelles » du geste dans la manipulation tactile d'objets graphiques en ce sens où le geste ergotique engage des formes complexes de la perception sensori-motrice qui sont encore inexploitées par les interfaces multi-tactiles. L'absence de feedback kinesthésique et la faible granularité du retour tactile, quand bien même ce dernier serait absolument synchrone (aucune latence), écartent considérablement les interactions multi-tactiles des caractères naturels tels que nous les avons définis préalablement. Malgré l'affordance perçue des objets conférée par le réalisme de leur représentation graphique et le caractère réaliste de leur comportement régi par des modèles de physique naïve, l'absence de préhension des objets, imposée par la morphologie de ces interfaces, nous éloigne considérablement du paradigme d'interface « naturelle ».

La dichotomie entre l'affordance perçue et l'affordance réelle des objets graphiques suggère les limites de la métaphore du réel dans le couplage homme-machine, dans la rencontre entre l'espace réel et l'espace virtuel. Selon Emmanuel Kant, l'espace serait une condition *a priori* de l'expérience. Il faut que l'espace fasse l'épreuve de notre subjectivité et se donne à notre sensibilité afin qu'il soit possible que nous en ayons l'intuition. Dans cette course au réalisme, l'« espace virtuel » devient un lieu de représentation et de prolongement du réel. Les objets graphiques se parent de relief et s'inscrivent dans des espaces de simulation qui déjouent nos sens. Cet art de l'imitation est guidé, comme nous l'avons relevé au cours de cette étude, par le postulat d'usabilité qui rapproche la simplicité d'utilisation d'une interface à sa capacité à proposer une relation impliquant des compétences que l'utilisateur a déjà acquises au cours de sa confrontation avec le réel. Toutefois, en donnant à l'utilisateur les moyens de toucher cette illusion des doigts, je crains que la supercherie, une

fois découverte par les sens haptiques, n'inscrive le réalisme des interactions davantage dans un art du simulacre que dans celui de la simulation.

Au cours de la phrase gestuelle surgit un moment où le décalage entre la faible qualité du feedback tactilo-kinesthésique proposé par les périphériques multi-tactiles et le réalisme excessif de la mimésis visuelle révèle à l'utilisateur que les règles d'interaction naturelle ne sont pas respectées. C'est l'exploration même de l'interface, la découverte, l'heuristique d'interaction qui est altérée puisque sa relation avec l'utilisateur repose sur un « mensonge », une illusion haptique, un « vertige sensitif ». La proprioception, c'est-à-dire la capacité à appréhender l'espace par le biais de nos sens, est alors sans cesse mise en déroute par l'univers numérique qui évolue de façon parfois « logique », en reproduisant de manière « naïve » la représentation et le comportement des objets, et de façon parfois contradictoire en contrariant la perception synesthésique. J'emprunterai le terme « propriodéception » proposé par Derrick de Kerckhove²⁰³ pour qualifier cette déconvenue ou « dés-illusion » éprouvée par l'utilisateur. Celui-ci prend alors conscience que les règles d'interaction dont il fait l'expérience sont faussées ou en quelque sortes truquées, car elles reposent sur une contrefaçon incomplète du réel.

A travers cette étude, je souligne également les limites de l'implication des acquis de la communication gestuelle dans l'« intuitivité » des interactions. D'un côté du continuum de Kendon, le faible degré de conventionnalité du geste co-verbal, sa nature spontanée et protéiforme, posent le problème de sa participation à la création d'un modèle d'interaction. De l'autre côté du continuum, les emblèmes, ayant pourtant un fort degré de conventionnalité, ne sont pas pour autant universels puisqu'ils dépendent dans une certaine mesure du contexte socio-culturel de l'individu. L'abstraction gestuelle effectuée par l'utilisateur de manière irrégulière et spontanée est imposée par le degré d'intégration extraordinairement faible des périphériques tactiles et répond alors davantage aux règles de l'exploration de l'interface, de l'expérience des interactions et de leur apprentissage, qu'à la reproduction de gestes de la communication humaine. Ce constat pose donc la question de la définition d'une grammaire gestuelle, de son adoption par les utilisateurs et des enjeux de sa standardisation. C'est dans le souci d'accompagner ces réflexions par la pratique et la

²⁰³ KERCKHOVE Derrick (de). *Propriodéception et Autonomation*. In *Les cinq sens de la création : art, technologie, sensorialité*. Champ Vallon. Seyssel, 1996, pp.132-140 (Coll. Milieux)

nécessité d'appliquer une démarche empirique, que j'ai envisagé de concevoir un environnement de développement (*framework*) multi-tactile.

Approche pratique

L'élaboration de prototypes matériels répondait aux besoins formulés par l'entreprise Virtual-IT, partenaire du projet, d'orienter son activité vers le design d'interaction multi-tactile. Déclinés selon différentes configurations, ces dispositifs nous ont permis de disposer d'un cadre d'expérimentation et d'observation idéal pour valider et enrichir notre modèle d'interaction. La fabrication d'une borne et d'une table interactive nous a permis de participer à plusieurs événements professionnels, comme *Laval Virtual* et *Euronaval*, ou d'autres tout public comme le *Toulouse Game Show* afin de soumettre nos développements logiciels à différentes catégories socioprofessionnelles. En effectuant l'enregistrement de gestes utilisateur par l'intermédiaire d'un logiciel d'acquisition de points de contact développé en ce sens, j'ai été en mesure de valider et d'enrichir les principes décrits par le modèle d'interaction. L'analyse des enregistrements m'a également permis d'ajouter aux limites conceptuelles de la métaphore du réel le phénomène d'ambiguïté d'échelle entre la dynamique du geste utilisateur et la taille perçue des objets graphiques. Par ailleurs, l'étude de la phrase gestuelle a révélé sa nature protéiforme, en termes de configuration et de « dynamicité », et sa nature évolutive, en termes de points de contact de référence. Enfin, il m'a été possible de comprendre en quoi les technologies d'acquisition (sensibilité optique, configuration) intervenaient dans la détermination du début et de la fin de la phase nucléaire. Les retours d'expérience utilisateur, les observations faites sur le terrain et l'analyse des enregistrements ont participé à la constitution d'un espace de conception et à l'enrichissement du modèle d'interaction sur lesquels a reposé le développement de l'environnement multi-tactile.

Les limites des API multi-tactiles existantes nous ont en effet imposé de concevoir notre propre environnement de développement multi-tactile. Celui-ci devait répondre à plusieurs objectifs :

- D'une part, en inscrivant ce projet dans un cadre professionnel et commercial et dans une logique de production et de pérennisation des développements, il devait nous permettre d'unifier le développement d'applications multi-tactiles. Pour y parvenir,

nous nous sommes affranchis de la dépendance des applications au périphérique multi-tactile en créant un format générique de description de points de contact, en assurant la compatibilité avec les protocoles existants (TUIO, TCP Flash binary, XML socket) et en développant une architecture logicielle autour d'un intergiciel et d'interfaces de développement (API).

- Cet environnement de développement devait également faciliter la réalisation d'applications multi-tactiles et réduire les temps de production ou d'intégration. Pour cela, nous avons déplacé conceptuellement l'ensemble des fonctions de traitement et d'analyse des points de contact, généralement situé côté application, vers l'intergiciel. En organisant notre environnement de développement autour d'un système « original » de communication événementielle entre l'intergiciel et les API, nous avons considérablement simplifié l'approche du développement d'applications multi-tactiles pour un développeur non spécialisé.
- Enfin, cet environnement de développement devait profiter des recherches et des développements antérieurs en intégrant les principes définis par le modèle d'interaction. C'est la raison pour laquelle nous avons adapté la classification OCGM (Objets, Containers, Gestes et Manipulations) proposée par George et Blake à la structure et intégré les caractères des interactions tactiles à la définition des patrons d'interaction. Nous avons également pris en compte les précisions amenées par l'analyse des enregistrements effectués sur le terrain en implémentant un algorithme de lissage et de correction de points de référence, puis nous avons ajouté des propriétés supplémentaires aux patrons d'interaction comme l'encadrement du nombre de points de contact et le delta d'activation.

C'est lors des phases de rencontre avec le public que nous avons fait l'observation du « vertige sensitif » évoqué précédemment. La rupture de la perception sensori-motrice conduisait l'utilisateur à réaliser que les modalités de communication avec la machine ne reflétaient pas ceux dont il avait l'intuition, nous observions alors la frustration vécue par l'utilisateur par son agacement ou son abandon. Alors qu'il était trop tôt à ce moment de la recherche pour saisir toute la complexité du phénomène, nous faisons toutefois le constat que les utilisateurs, bien que montrant une certaine aisance avec les interactions de

manipulation traditionnelles, manifestaient des signes d'incompréhension et de frustration dès que l'interface ne réagissait pas comme ils le supposaient. Lorsque le degré de conventionnalité nécessaire à certaines interactions était trop élevé et répondait davantage au parti pris du designer d'interaction, l'utilisateur se trouvait contraint à parler un langage qui lui était inconnu et cela contrariait l'idée suggérée à l'utilisateur d'une interface « naturelle ». Lorsque les modalités de la manipulation d'objets graphiques dans l'espace en deux dimensions de l'interface s'écartaient trop de la manipulation d'objets réels, l'utilisateur ne disposait pas suffisamment de retour visuel et haptique pour comprendre l'origine de ses erreurs et procéder efficacement à l'apprentissage des règles d'interaction.

La modélisation temporelle du geste, à travers sa forme canonique et son couplage avec le système tactile, et la segmentation de la phrase gestuelle nous ont permis de révéler le caractère intentionnel du geste à travers sa phase nucléaire. Par sa modélisation spatiale, nous avons relevé des ensembles finis de configuration et de « dynamicité » possibles qui ont contribué à la classification du geste en situation d'interactions multi-tactiles. En nous penchant sur les modalités d'interaction de ces interfaces, nous avons relevé l'influence positive de leur faible degré d'indirection sur le feedback visuel et l'implication de leur fort degré de compatibilité dans la réduction de la charge cognitive liée à la réalisation d'une tâche. Cependant, le degré d'intégration extrêmement faible de ces périphériques, résultat du nombre disproportionné de configurations et de degrés de liberté de la main par rapport à l'interface, semble nuire profondément à l'« intuitivité » des interactions, à la reconnaissance d'erreurs de manipulation et donc à leur apprentissage. Les résultats de cette étude indiquent que si les utilisateurs ne reproduisent pas spontanément les gestes de manipulation et de communication « naturels », c'est qu'ils sont d'une part contraints d'adapter leurs gestes, de leur forme spatiale initiale, en trois dimensions, vers une forme limitée aux deux dimensions de la surface, et d'autre part, que cette transformation implique un fort degré de conventionnalité nuisant ainsi à la découverte des interactions.

Ouverture théorique

De récentes études²⁰⁴ ont démontré que le nombre d'erreurs dans la saisie de texte sur un téléphone portable tactile était considérablement réduit si l'écran disposait d'un retour haptique. Des recherches²⁰⁵ conduites par *Motorola* ont également précisé que 80% des sujets de leur expérience préféraient lorsque le feedback sonore était accompagné d'un feedback haptique. Les méthodes pour offrir un retour haptique à l'utilisateur et augmenter le réalisme des interactions sont actuellement en cours de développement. Sur les téléphones mobiles par exemple, pour remplacer les actionneurs inertiels fournissant un retour haptique par vibration, les fabricants travaillent à intégrer aux écrans tactiles des actionneurs haptiques (piézoélectriques ou de surface) permettant de déformer légèrement l'écran de manière dynamique là où apparaissent les objets graphiques à manipuler, pour simuler la présence physique de ces objets. Suite aux conclusions de ce travail de recherche, je ne peux qu'émettre quelques réserves sur les directions actuelles de la recherche et de l'industrie.

Dans cette course au réalisme, à l'hyperréalité, le design d'interaction suit la direction d'une représentation toujours plus convaincante du réel, en augmentant l'illusion optique par l'illusion haptique. Le photo-réalisme, les moteurs 3D, la physique naïve et l'intelligence artificielle suffisaient pour tromper nos sens et nous faire vivre l'hallucination visuelle d'une réalité singulière. Toutefois le sens du toucher vient généralement confirmer ou infirmer les réalités perçues par la vue car il vient généralement dans un deuxième temps, capable d'explorer et d'apprécier les formes, les textures et les forces d'un objet. Le toucher est le dernier rempart de notre incrédulité, le sens qui accède aux propriétés invisibles de la matière, restées sous silence, cachées dans l'ombre ou hors de notre champ de vision. Au cours de l'exploration d'une interface tactile, toucher l'image revient inconsciemment à éprouver l'authenticité de la réalité qu'elle expose. C'est alors que certaines illusions vacillent car elles ne reposent que sur la pâle copie du réel, sur un « faux » au sens que lui prête Umberto Eco, dont la supercherie est si grande qu'elle finit quasiment toujours par être démasquée. Je crois qu'il est réducteur de considérer que l'espace virtuel ne puisse

²⁰⁴ BREWSTER Stephen, CHOCHAN Faraz, et BROWN Lorna. *Tactile feedback for mobile interactions*. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (CHI '07). ACM, New York, USA, 2007, pp. 159-162

²⁰⁵ CHANG Angela et O'SULLIVAN Conor. Audio-haptic feedback in mobile phones. In *CHI '05 extended abstracts on Human factors in computing systems* (CHI EA '05). ACM, New York, USA, 2005, pp. 1264-1267

exister que sous la forme d'une simulation du réel et non sous une forme de réalité singulière puisque cela reviendrait à ignorer les spécificités même de l'image numérique. En effet, la simulation numérique, algorithmique, parle son propre langage et s'adresse à nos sens selon des mécanismes complexes d'interaction qui inscrivent les objets de la simulation dans un univers « tangible », par les sens comme par l'esprit. Créer des univers virtuels dont l'espace d'activation ne serait régi que par les lois spatio-graphiques et physiques du réel reviendrait à soumettre de manière paradoxale les technologies numériques aux contraintes auxquelles elles sont pourtant censées pouvoir échapper. Comment pourrions-nous alors représenter et activer certaines actions abstraites du système sans entrer en conflit avec le paradigme réaliste ? Je suis convaincu que les technologies numériques, dont la capacité à faire coexister virtuellement des réalités contradictoires n'est plus à démontrer, supposent obligatoirement une forme d'abstraction, d'adaptation et d'apprentissage de la part de l'utilisateur. Sans tenir compte de ces spécificités, le paradigme d'interface utilisateur naturelle irait à l'encontre de la nature protéiforme des espaces et des temps virtuels, autonomes, extensibles, rétractables, multipliables et réglés par des logiques réalistes comme paradoxales. En suivant la définition proposée par Philippe Quéau²⁰⁶, ces « *hyper-images* » comporteraient plusieurs niveaux de réalité et de simulation, juxtaposés, corrélés ou fusionnés.

Ne devrions-nous pas déplacer l'affectation du terme « naturel », de la description des modalités de la relation homme-machine à celle des caractères singuliers de ces hyper-images. Il nous serait alors possible d'aborder le design d'interaction en intégrant le nouvel ordre de représentation proposé par ces images, mêlant en permanence le vrai et le faux, le synthétique et le naturel, le réel et le virtuel. Le virtuel deviendrait aussi réel que le réel, voire même plus réel. En créant les conditions pour que ces images encouragent nos sens à éprouver des sensations tactiles et kinesthésiques nouvelles, des distorsions inhabituelles de la perception et des modes dialogiques spécifiques, le design d'interaction s'inscrirait davantage dans les formes de la relation proposées par ces images. Contrairement à la définition classique des interfaces utilisateur naturelles, qui suppose de manière assez réductrice d'adapter les interfaces aux compétences intellectuelles d'un nourrisson, je défends l'idée qu'il devient nécessaire de nous interroger sur nos limites physiques et psychologiques et sur notre faculté d'intégrer ces objets technologiques dans les processus

²⁰⁶ QUÉAU Philippe. *Metaxu – Théorie de l'art intermédiaire*. Champ Vallon. Seyssel, 1989, 344 p. (coll. Milieux)

naturels de l'apprentissage comme nous le faisons déjà pour des objets techniques plus ordinaires. Car ces hyper-images appartiennent désormais au paysage contemporain, non comme des artefacts de science-fiction mais davantage comme des objets usuels. Les nouvelles générations apprennent aujourd'hui à se servir des technologies informatiques dès leur plus jeune âge et intègrent les modalités de leur utilisation comme une compétence que nous pourrions qualifier alors de « naturelle ».

Travaux futurs

Nous disposons aujourd'hui d'un environnement de développement stable et particulièrement ouvert à de futures directions de recherche. Au cours du développement des interfaces de programmation (API), nous avons mis à l'épreuve les principes de communication asynchrones de données tactiles entre l'intergiciel et les applications client par l'intermédiaire des applicatifs développés en entreprise. Notre feuille de route prévoyait d'obtenir à l'issue de cette étude un environnement de développement intégrant les gestuelles et la majorité des interactions de manipulation couramment utilisées dans ce type d'interfaces afin de valider notre modèle théorique. Nous sommes désormais équipés pour poursuivre les recherches. Nous envisageons de nous appuyer sur ces développements préliminaires pour implémenter de nouveaux patrons d'interaction et orienter nos efforts vers le design d'interfaces.

Au cours des phases d'observation, nous avons remarqué la présence de points fantômes et de points aberrants issus des phénomènes d'occlusion qui intervenaient au cours des interactions. Avec l'évolution des technologies multi-tactiles, la présence de points fantômes a quasiment disparu. L'implémentation d'un algorithme de correction des points de référence nous a permis de supprimer l'effet des effleurements involontaires et des changements intervenant au cours de la phrase gestuelle. Toutefois, l'utilisation de filtres de lissage demeure indispensable pour corriger l'irrégularité de certains gestes et l'instabilité de la pose statique détectée par certains capteurs. Nous avons choisi d'utiliser le lissage par moyenne mobile pondérée pour la simplicité de son implémentation et pour obtenir rapidement un environnement de développement fonctionnel. Le filtre Kalman étant souvent utilisé pour estimer les états d'un système dynamique à partir de données incomplètes ou bruitées, nous prévoyons d'implémenter cet algorithme pour lisser les

valeurs des événements `InteractionEvent`. Nous pensons qu'en utilisant un filtre de type prédictif, nous serons en mesure de réduire les temps de latence ressentis par l'utilisateur. Enfin, nous envisageons d'effectuer des mesures précises de la latence générée par ces différents filtres et de manière générale, diminuer la latence globale de notre environnement logiciel et matériel. Par l'intermédiaire d'observations et de questionnaires, nous souhaitons pouvoir évaluer le degré d'« immédiateté » ressenti par l'utilisateur lors des interactions de manipulation afin d'obtenir un temps de latence maximum théorique.

Au cours de cette étude, nous avons souligné les difficultés rencontrées par les utilisateurs dans l'exploration et l'apprentissage des interactions multi-tactiles ayant un fort degré de conventionnalité. Parmi les directions de recherche, nous envisageons d'intégrer à notre environnement de développement un système d'apprentissage didactique dont la fonction sera d'accompagner l'utilisateur dans la découverte des interactions définies par les designers et les développeurs, et d'en faciliter ainsi l'apprentissage. Dans un premier temps, nous souhaitons que l'intergiciel soit en mesure de détecter les erreurs de manipulation de l'utilisateur, comme la réalisation d'une gestuelle non reconnue ou réalisée incorrectement, puis d'envoyer à l'API client cette information pour que l'application fournisse à l'utilisateur une assistance sous la forme d'un feedback visuel représentant le ou les interactions attendues par l'objet ou le container concerné. Nous envisageons par la suite de faire évoluer ce système en tenant compte des possibilités techniques permises par l'évolution des technologies tactiles. Nous souhaiterions profiter, en plus des informations de contact, des informations liées à la position, à la distance et à l'angle d'approche des mains, afin de proposer une assistance à l'exploration de l'interface en affichant les interactions permises par l'objet ou le container visé par le geste et ce tout au long de la phase d'approche de la phrase gestuelle. Cette assistance visuelle disparaîtrait aussitôt que l'utilisateur toucherait l'objet et pourrait réapparaître plus tard en fonction des capacités d'apprentissage de l'utilisateur. Il serait également possible que l'assistance visuelle intervienne au cours même de la phase nucléaire du geste en affichant en temps réel les continuités possibles du geste en fonction des interactions prédéfinies.

Les futures orientations de notre recherche tiendront compte des observations et des conclusions de cette étude. Nous tenterons d'approcher le paradigme d'interface naturelle en intégrant davantage l'influence de la surface interactive, à la fois surface de contact et de projection, sur les modalités de la relation homme-machine. L'écran multi-tactile, sans

texture ni relief, sans retour de force ni élasticité, ne saurait procurer toute la granularité nécessaire aux objets virtuels pour garantir l'illusion. Il me semble nécessaire de considérer la surface de contact non plus comme un espace d'actualisation direct reposant sur l'illusion optique et haptique mais davantage comme un espace de médiation entre l'utilisateur et l'interface. Il s'agit de reconnaître la matérialité des dispositifs-images et d'approcher leur surface comme une zone d'interaction mixte et singulière, comme un espace d'entre-deux, d'in-ter-action. L'écran est parfois réfléchissant, comme un miroir sans teint en imitant les actions de l'utilisateur (manipulation) reproduites par un analogon numérique, parfois infléchissant, comme un miroir déformant en transformant les actions de l'utilisateur selon des règles d'interaction plus abstraites (gestuelle), et parfois agissant, comme une enveloppe, une « seconde peau », une matière invisible mais palpable, donnant à l'utilisateur une sensation de contrôle, presque démiurgique, de l'univers virtuel.

Les travaux de *Jean-Baptiste de la Rivière et al.*²⁰⁷, présentés lors du salon SIGGRAPH en 2010, vont justement dans ce sens. En réalisant que les images en trois dimensions s'adressaient davantage au sens de la vue et qu'en essayant de toucher ces images, l'illusion est trahie par la main passant au travers des objets graphiques, l'équipe de recherche et de développement de la société Immersion a mis au point un principe d'inscription de l'univers 3D dans la profondeur de la table. Le système consiste en une table interactive et des lunettes polarisées dont l'orientation est analysée en temps réel permettant à deux utilisateurs de voir une simulation en trois dimensions selon des angles de vue dépendants de leur position respective grâce à la synchronisation entre l'obturateur et la diffusion d'images. Au moment où la main approche de l'écran, le plan horizontal de la simulation 3D s'abaisse et l'utilisateur peut alors interagir avec l'univers sans avoir l'impression que sa main traverse les objets 3D.

Ce qui est particulièrement pertinent dans ce dispositif, c'est la manière dont les objets en trois dimensions s'inscrivent dans l'espace de la manipulation en restant à une certaine distance (virtuelle) de la main. La manipulation des objets 3D ne se réalise plus selon l'illusion de leur préhension mais plutôt selon une relation plus subtile, rendue évidente par la présence d'un retour visuel de lignes blanches verticales en trois dimensions reliant les

²⁰⁷ RIVIERE Jean-Baptiste (de la), KERVEGANT Cédric, DITTLO Nicolas, COURTOIS Mathieu et ORVAIN Emmanuel. *3D multitouch: when tactile tables meet immersive visualization technologies*. In *ACM SIGGRAPH 2010 Emerging Technologies* (SIGGRAPH '10). ACM, New York, USA, 2010, 1 p.

zones de contact à la surface des objets situés (virtuellement) en dessous (cf. figure 104). Ces lignes sont comme les baguettes dans un théâtre d'ombre, les fils dans un spectacle de marionnettes. Mais ici, c'est la main qui sert de croix, de *contrôle*, de manipulateur, en prolongement les doigts par des fils que l'on montre avec ostentation. La main s'inscrit dans les deux espaces, elle devient un instrument mixte disposant de propriétés physiques et virtuelles. Les fils sont rendus visibles, les ressorts de la manipulation sont montrés, car il s'agit davantage de manifester à l'utilisateur les mécanismes de la relation que de créer une illusion. Contrairement au marionnettiste qui garde jalousement ses secrets de fabrication, qui dissimule la manière dont sa marionnette est reliée au contrôle²⁰⁸ en recouvrant la croix à la fin de chaque représentation, le designer d'interaction provoque ici leur découverte par l'utilisateur. L'exploration de l'interface révèle l'ensecret²⁰⁹ du design d'interaction et suggère à l'utilisateur les artifices de son contrôle. La « manipulation » ne vise pas à « animer » un objet inerte. L'utilisateur manipule, joue et interagit avec l'interface selon ses règles, ses degrés de liberté et sa propre autonomie. Le « jeu » est garant de l'intuition de réalité en ce sens où le rapport de l'utilisateur aux objets virtuels est régi par des règles nouvelles et actualisé par la tension exercée par l'utilisateur sur les objets situés à l'autre l'extrémité. Au jeu correspond la dimension ouverte de l'interaction, comme un « *espace ménagé pour qu'un libre mouvement s'exerce*²¹⁰ », celui du geste et de la main.

²⁰⁸ Manipulateur permettant d'animer la marionnette.

²⁰⁹ L'Ensecrèt ou ensecrètement est le fait de relier la marionnette à son contrôle, ou croix, par des fils en équilibrant le fantoche afin qu'il marche correctement.

²¹⁰ BOISSIER Jean-Louis. *La relation comme forme - L'interactivité en art*. Mamco. Genève, 2004, 311 p.



Figure 104 - Inscription 3D de la main dans l'espace virtuel. Capture d'écran. « 3D Multitouch: When Tactile Tables Meet Immersive Visualization Technologies » In Youtube. Disponible ici : <http://youtu.be/dAdEpaxv3p4>

Bilan du CIFRE

Enfin, je rappellerai combien la collaboration soutenue entre les laboratoires universitaires, LARA et IRIT, et l'entreprise Virtual-IT, sous la forme d'un dispositif CIFRE, a été profitable. Je souhaiterais témoigner de la richesse de cette expérience et de la fécondité de la circulation des concepts entre la critique universitaire et le pragmatisme industriel. La confiance que m'ont accordée mes directeurs de recherche, monsieur Gilles Méthel et monsieur Jean-Pierre Jessel, en soutenant cette thèse et les responsabilités qui m'ont été confiées par les directeurs de Virtual-IT, monsieur Jean-François Pal et madame Cécile Dos Santo, m'ont projeté au cœur des processus industriels de l'innovation technique, répondant ainsi au désir que j'entretenais de me sentir impliqué dans l'évolution du design d'interaction. Ces conditions remarquables m'ont permis de mettre un pied dans le bassin industriel régional (partenaires et sous-traitants), d'approcher les acteurs de l'innovation (MPI, INPI, OSEO) et les écoles d'ingénieurs (école des Mines d'Albi, Epitech). La dynamique et l'investissement de tous ces intervenants nous ont permis de présenter ces travaux lors de salons technologiques (Laval Virtual, Imagina), lors de conférences (Ludovia, Epitech) ou dans le cadre de démonstrations internes. L'intérêt et la confiance que nous ont manifestés tout au long du projet des acteurs importants du secteur industriel (DCNS, Airbus) ou des institutions

publiques (Hôtels de Région) et culturelles (musées) nous encouragent aujourd'hui à poursuivre ce travail de recherche.

Animé par le besoin de vivre de « belles choses, ce projet n'aurait su tant me satisfaire si les hommes ayant contribué à sa genèse n'avaient été guidés par un « *souffle de folie* », celui qu'André Gide disait appartenir au monde des rêves et qu'il m'a pourtant été possible de percevoir.

ANNEXES

1. Annexes générales

1.1. Cycles d'innovation de quelques technologies

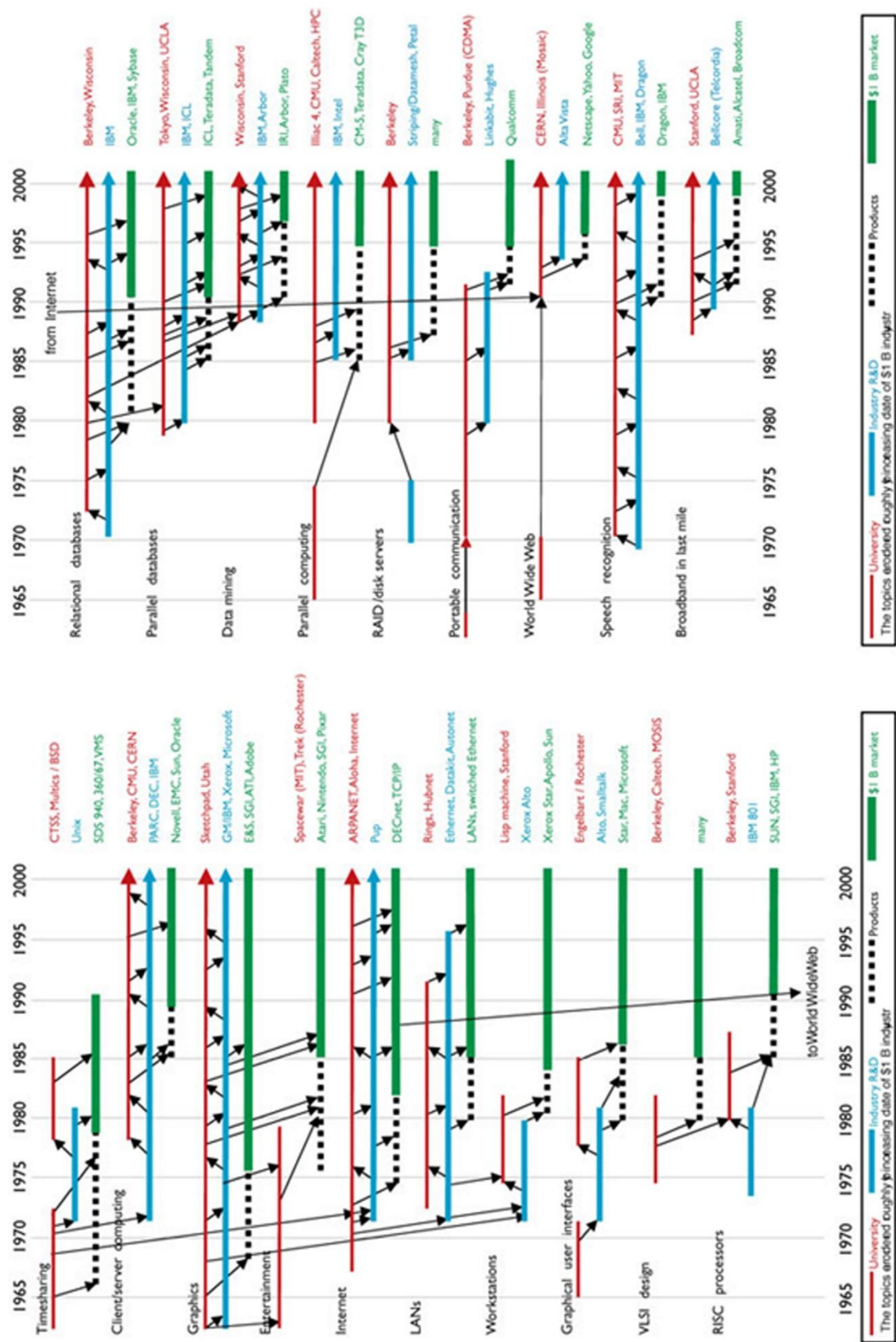


Figure 105 - Cycles de l'innovation concernant le domaine des TIC.

1.2. VideoHarp



Figure 106 - *VideoHarp*, produite par la société *SensorFrame*, développé initialement par Paul McAvinney et Dean Rubine.

1.3. Extrait complet de « 2001, l'Odyssée de l'espace » d'Arthur C. Clarke

« Lorsqu'il en eut assez des rapports officiels et des mémos, il brancha son minibloc d'information sur le circuit du vaisseau et parcourut les dernières nouvelles de la Terre. Il formait l'un après l'autre les numéros de code des principaux journaux électroniques du monde. Il en connaissait par coeur la plupart et n'avait pas besoin de consulter la liste qui figurait au dos du bloc. En jouant sur la mémoire de la visionneuse, il pouvait consulter la première page et choisir rapidement les rubriques qui l'intéressaient. Chacune avait son propre numéro de référence et, lorsqu'il le formait, le rectangle qui avait les dimensions d'un timbre-poste s'agrandissait sur l'écran. La lecture achevée, il suffisait de revenir à la vision de la page entière et de choisir une autre rubrique.

Parfois, Floyd se demandait si le minibloc et la technologie fantastique qu'il supposait représentaient le sommet des découvertes humaines en matière de communications. Il se trouvait en plein espace, s'éloignant de la Terre à des milliers de milles à l'heure et pourtant, en quelques fractions de seconde, il lui était possible de consulter n'importe quel journal. Le mot même de journal était une survivance anachronique en cet âge électronique. Le texte se modifiait automatiquement d'heure en heure. Même en ne lisant que la version anglaise on pouvait passer sa vie entière à absorber le flot sans cesse changeant des informations retransmises par satellites.

Il était difficile d'imaginer que le système pût être modifié ou amélioré. Pourtant, songea Floyd, tôt ou tard il disparaîtrait pour être remplacé par quelque chose qui renverrait les miniblocs au rang des presses de Gutenberg.

La lecture des journaux électroniques amenait souvent une autre réflexion : plus les moyens de diffusion se faisaient merveilleux, plus barbare, atterrant et choquant était leur contenu. Accidents, désastres, crimes, menaces de conflit, éditoriaux sinistres - tels semblaient être les sujets principaux des articles qui se propageaient dans l'espace. Floyd en venait parfois à se demander si tout cela était vraiment aussi terrible qu'il y semblait. Les informations d'Utopie, après tout, auraient sans doute été atrocement ennuyeuses. »

2. Design et concepts art

2.1. Prototype « inTouch »

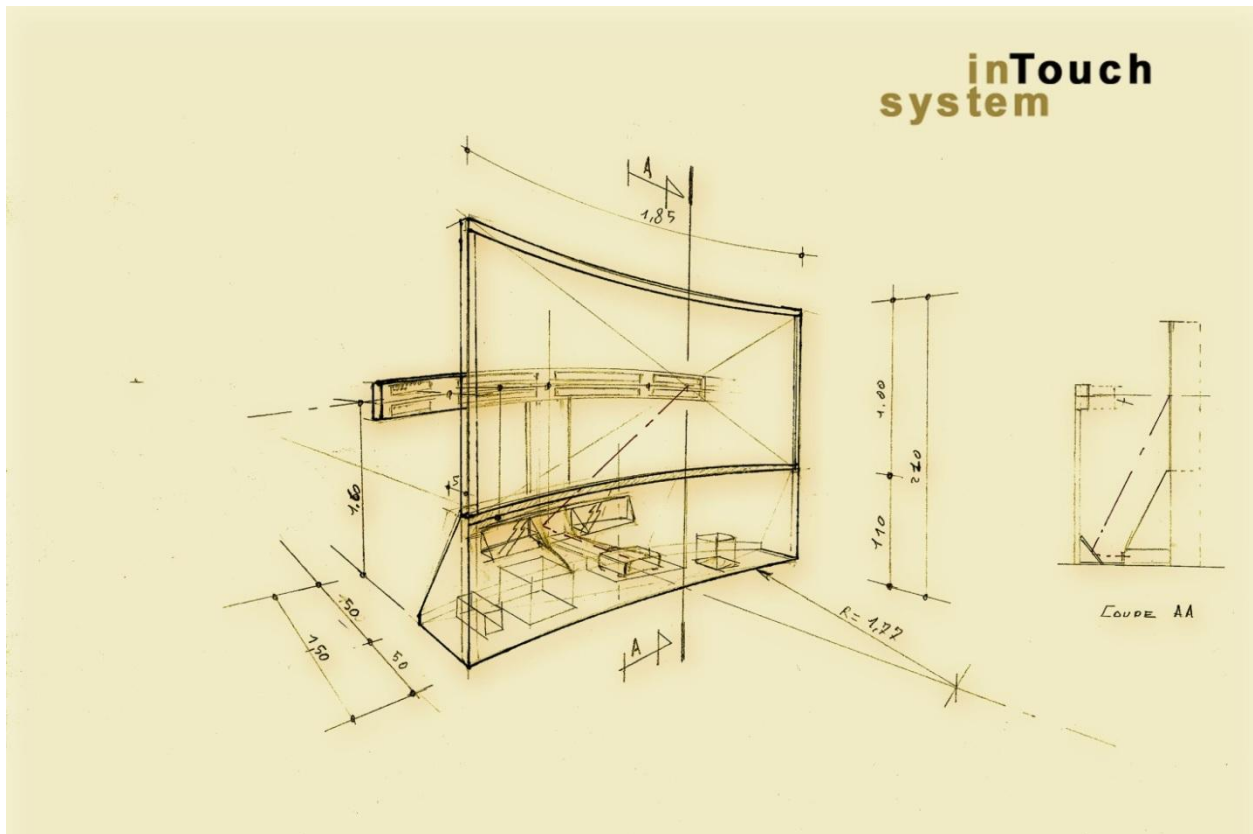


Figure 107 - Concept art d'avant-projet. Croquis réalisé par Jean-François Pal.

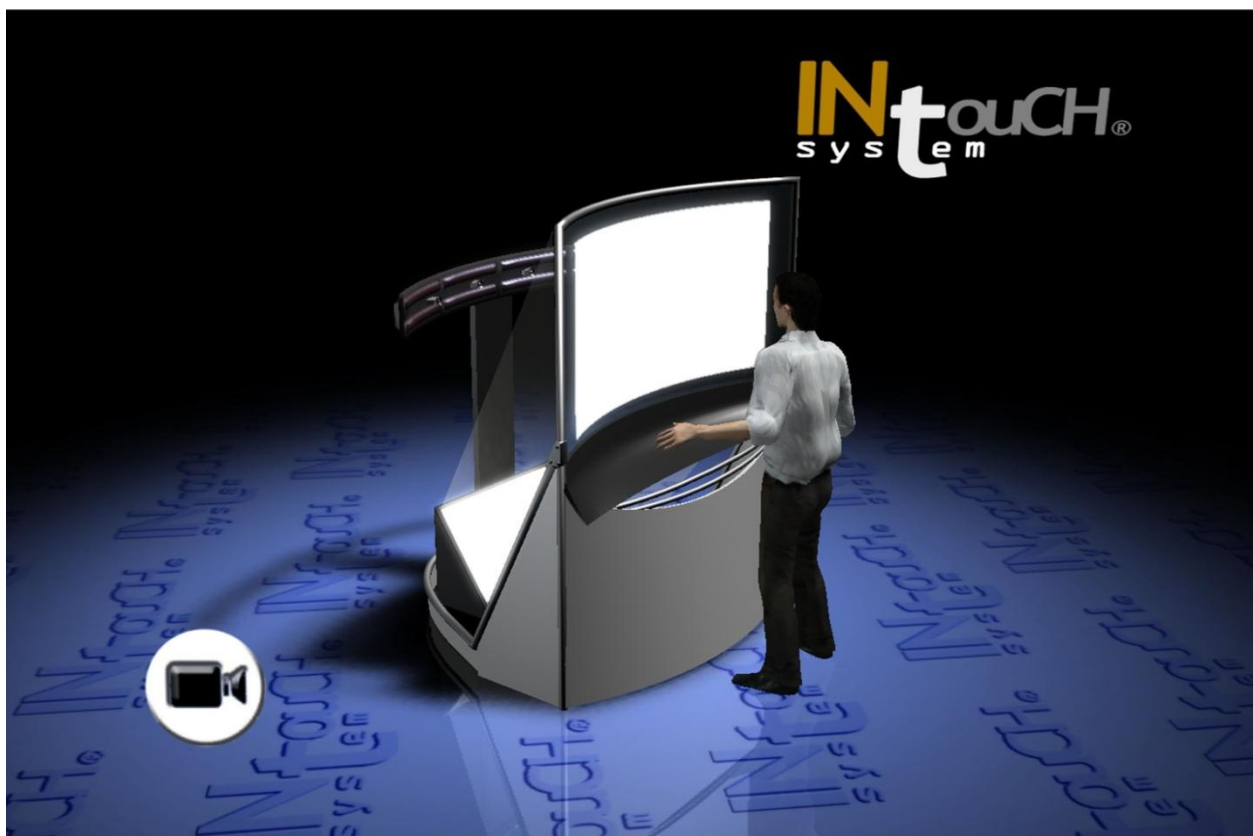


Figure 108 - Scène 3D issue du concept art.

2.2. Prototypes divers



Figure 109 - Concept art d'un prototype d'écrans multi-tactiles circulaires translucides.



Figure 110 - Concept art d'un prototype de réalité augmentée.

2.3. Projet « LUNAR »

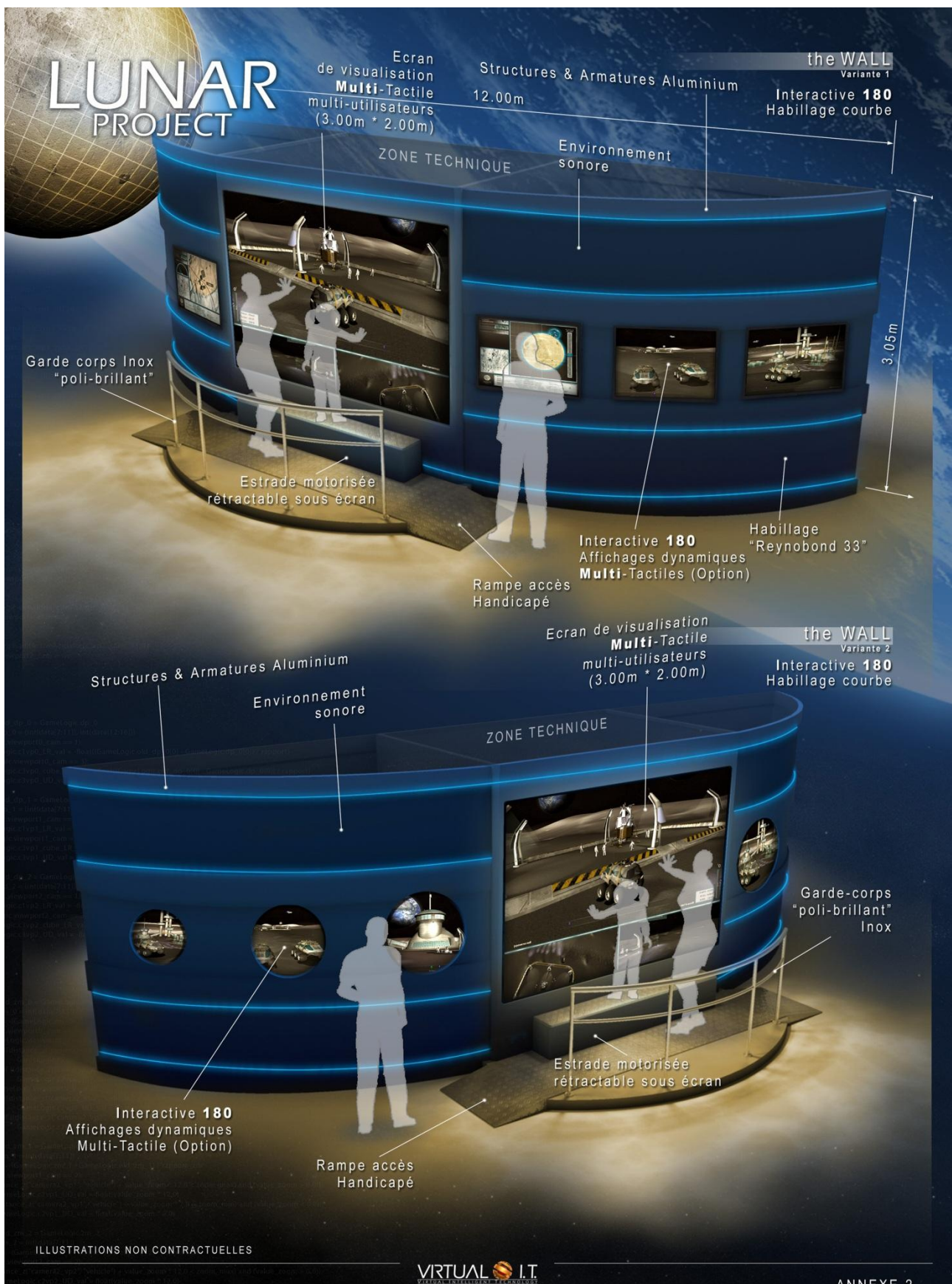


Figure 111 - Extrait de la plaquette de présentation du projet LUNAR.

2.4. Projet « QUARIUS »

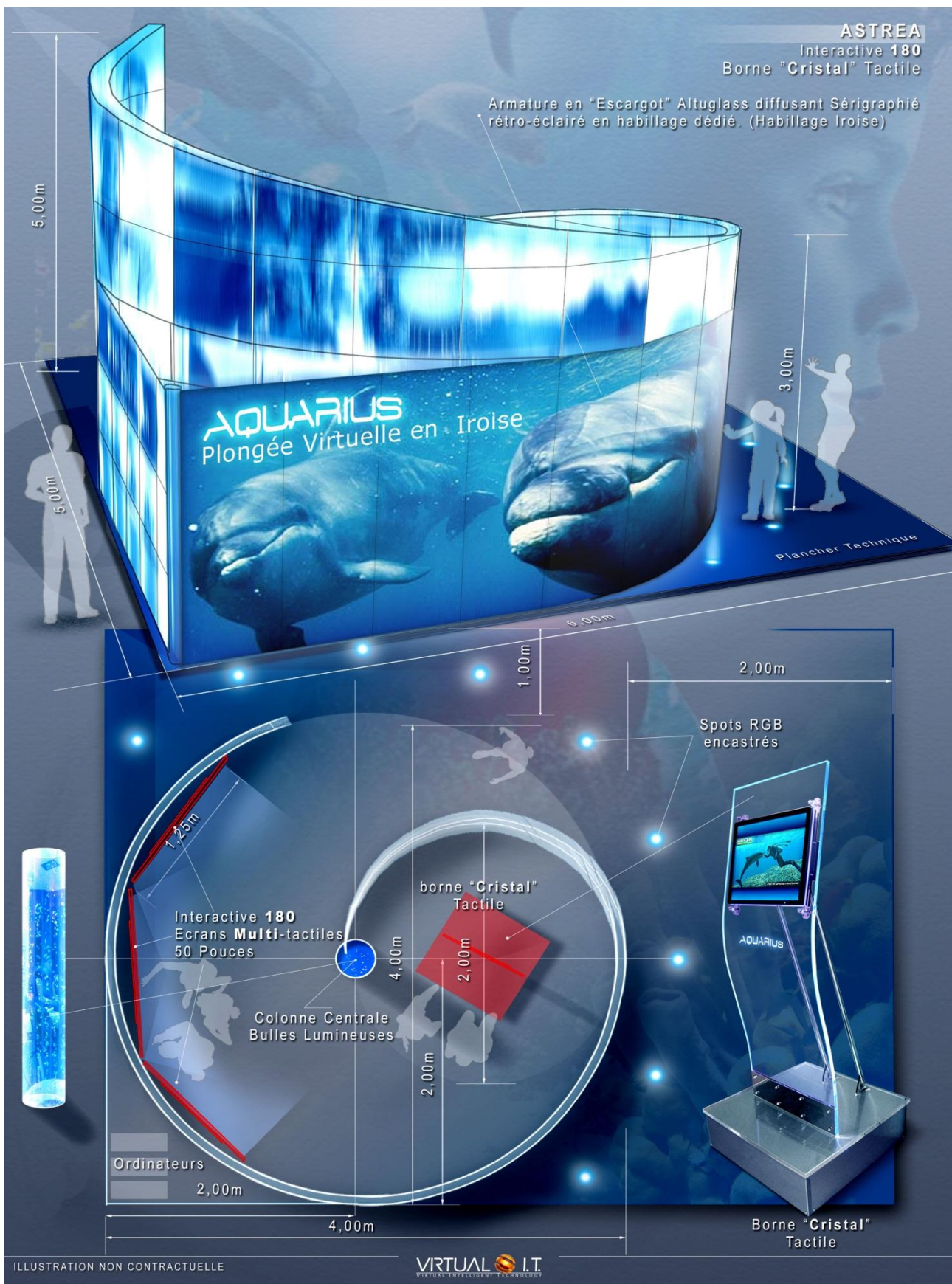


Figure 112 - Extrait de la plaquette de présentation du projet AQUARIUS.

2.5. Salon DCNS - Toulon, le Mourillon



Figure 113 - Photo du showroom réalisé pour DCNS, site de Toulon.

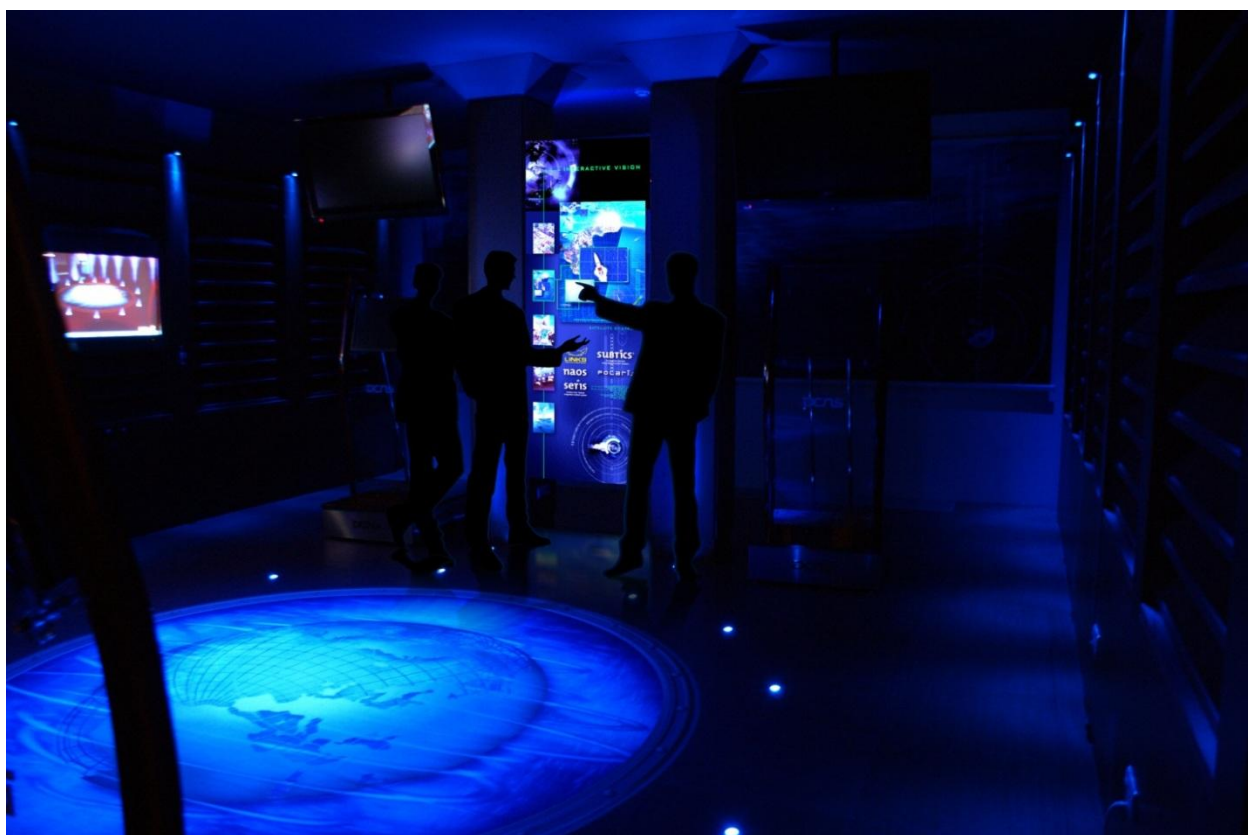


Figure 114 - Photo du showroom réalisé pour DCNS, site de Toulon.

3. Salons et démonstrations

3.1. DCNS - Démonstration interne – Toulon, 2009



Figure 115 - Vue d'ensemble de la démonstration réalisée devant le pôle communication et R&D de DCNS Toulon



Figure 116 - Vue d'ensemble de la démonstration réalisée devant le pôle communication et R&D de DCNS Toulon

3.2. Milipol - Paris, 2009



Figure 117 - Vue d'ensemble du showroom DCNS au salon Milipol 2009.



Figure 118 - Simulation 3D multi-tactile présentée lors du salon Milipol 2009.

3.3. Toulouse Game Show – Toulouse, 2009



Figure 119 - Un père apprenant à son fils comment utiliser l'interface du projet LUNAR.



Figure 120 - Un utilisateur en train de jouer au jeu Warcraft III adapté pour le multi-tactile.

3.4. ARICAD - Paris, 2010



Figure 121 - Simulation 3D multi-tactile présentée lors du salon ARICAD.



Figure 122 - Conférence réalisée lors du salon ARICAD.

3.5. Laval Virtual - Laval, 2010



Figure 123 - Photo prise durant le salon Laval Virtual.



Figure 124 - Photo prise durant le salon Laval Virtual.

BIBLIOGRAPHIE

Ouvrages généraux

- ARENDT Hannah.** *Condition de l'homme moderne*. Calmann-Lévy. Paris, Pocket, 1983, 406 p.
- ARISTOTE.** *Les Parties des animaux*. Belles Lettres, 2003, 193 p. (Coll. Universités de France)
- BATAILLE Georges.** *Informe*. In *Œuvres complètes I*, Denis Hollier. Paris, Gallimard, 1970, 464 p. (Coll. blanche)
- BERGSON Henri.** *L'évolution créatrice*. PUF. Paris, 2003 (10e édition), 384 p. (coll. Quadrige, n°1)
- BOISSIER Jean-Louis.** *La relation comme forme - L'interactivité en art*. Mamco. Genève, 2004, 311 p.
- BULWER John.** *Chirologia or The Natural Language of the Hand*. Kessinger Publishing Co, 2003, 380 p.
- CADOZ Claude.** *Les réalités virtuelles*. Paris, Flammarion, 1994, Collection Dominos, 125 p.
- CLARKE Arthur C.** *2001 : L'Odyssée de l'espace*. Traduction : DEMUTH Michel. J'ai lu, 2001, 190 p.
- COUCHOT Edmond.** *Des changements dans la hiérarchie du sensible*. In *Les cinq sens de la création*, BORILLO Mario et SAUVAGEOT Anne. Champ Vallon, novembre 1998, 220 p. (Coll. Milieux)
- COUCHOT Edmond et HILAIRE Norbert.** *L'art numérique*. Flammarion, Paris, 2005, 265 p. (coll. Art)
- DICK Philip K.** *La trilogie divine, Tome 1 : Siva*. Editions Gallimard 2006, 353 p. (Coll. Folio SF)
- ELLUL Jacques.** *La technique, ou, L'enjeu du siècle*. Economica, 1999, 423 p.
- FOREST Fred.** *Pour un art actuel - L'art à l'heure d'internet*. L'Harmattan. Paris, 1997, 282 p. (coll. L'Art en bref)
- GIBSON James J.** *L'approche écologique de la perception visuelle*. MF, 2010, 332p. (Coll. Dehors)
- GIBSON James. J.** *The Senses Considered as Perceptual Systems*. Greenwood Press Reprint, 1983, 335 p.
- Gide André.** *Journal 1889-1939*. Gallimard, nouvelle édition, 1977, 1378 p. (Coll. Bibliothèque de la pléiade)
- GILLE Bertrand** (Ouvrage collectif sous la direction de). *Histoire des techniques : Technique et civilisations, technique et sciences*. Paris, Gallimard, 1978, 1652 p. (Coll. Encyclopédie de La Pléiade)
- HATWELL Yvette.** *Toucher l'espace. La main et la perception tactile de l'espace*. Lille, Presses Universitaires de Lille. 1986, 374 p.
- HATWELL Yvette, STRERI A. et GENTAZ E.** *Toucher pour connaître. Psychologie cognitive de la perception tactile manuelle*. Paris, PUF, 2000, 332 p.
- KERCHKHOVE Derrick (de).** *Propriodéception et Autonomation*. In *Les cinq sens de la création : art, technologie, sensorialité*. Champ Vallon. Seyssel, 1996, pp.132-140 (Coll. Milieux)
- KISSELEVA Olga.** *Cyberart, un essai sur l'art du dialogue*. L'Harmattan. Paris, 1998, 368 p. (coll. L'Ouverture Philosophique)
- LEROI-GOURHAN André.** *Le geste et la parole*. Paris, Albin Michel, 1964, 2 vol., 323 et 285 p.
- LÉVI-STRAUSS Claude.** *La pensée sauvage*. Omnibus, 8 janvier 1993, 393 p.
- MAEDA John.** *The Laws of Simplicity (Simplicity : Design, Technology, Business, Life)*. MIT Press, 2006, 127 p.
- McLUHAN Marshal.** *Understanding Media: The Extensions of Man*. MIT Press, 1994, 392 p.

McNEILL David. *Hand and Mind : What Gestures Reveal about Thought*. University Of Chicago Press, 1996, 423 p.

MORIN Edgar. *La Tête bien faite : Penser la réforme, reformer la pensée*. Seuil, 1999, 153 p. (Coll. L'histoire immédiate)

NIELSEN Jakob. *Usability inspection methods*. Morgan John Wiley & Sons, 1994, 448 p.

NORMAN Donald Arthur. *The Design of Everyday Things*. Basic Books, 2002, 272 p.

NORMAN Donald Arthur. *The Invisible Computer : why good products can fail, the personal computer is so complex, and information appliances are the solution*. MIT Press, Cambridge, septembre 1998, 316 p.

PIAGET Jean. *L'épistémologie génétique*. PUF, 2005, 126 p. (Coll. Que sais-je ?)

PIAGET Jean. *Psychologie et épistémologie*, Paris, Denoël, 1970, 85 p.

POINCARÉ Henri. *La science et l'hypothèse*. Paris, Flammarion, 1968, 252 p. (coll. Champs Flammarion)

QUÉAU Philippe. *Metaxu – Théorie de l'art intermédiaire*. Champ Vallon. Seyssel, 1989, 344 p. (coll. Milieux)

QUINTILIEN. *Institution oratoire*. Traduction : Jean Cousin. Paris. Belles Lettres. 1979, (Coll. Universités de France)

RASKIN Jef. *The Humane Interface : New directions for designing interactive systems*. Addison-Wesley Professional, 2000, 256 p.

RÉVÉSZ Géza. *Psychology and Art of the Blind*. London, Longmans Green, 1950, 338 p.

SCHMITT Jean-Claude. *La raison des gestes dans l'Occident médiéval*. Paris, Gallimard, 1990, 432 p.

SIMONDON Gilbert. *Du mode d'existence des objets techniques*. Aubier, Paris, 2001, 333 p.

STIEGLER Bernard. *La technique et le temps - La faute d'Épiméthée* (Tome 1). Galilée. Paris, 1994, 284 p.

WEISSBERG Jean-Louis, *Présences à distance - déplacement virtuel et réseaux numériques - Pourquoi nous ne croyons plus la télévision ?*. L'Harmattan. Paris, 1999, 304 p. (coll. Communication et Civilisation)

Revue scientifique - Actes de conférence

BAECKER R. M. *Towards a characterization of graphical interaction*. In *Human-computer interaction*. Morgan Kaufmann Publishers Inc., San Francisco, USA, 1987, pp. 471-482

BAECKER R. M. et **BUXTON William.** EDS. *Readings in Human-Computer Interaction : A Multidisciplinary Approach*. Kaufmann, Los Altos, Calif., 1987, pp. 357-365

BEAUDOUIN-LAFON Michel. *Instrumental interaction : an interaction model for designing post-WIMP user interfaces*. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (CHI '00). ACM, New York, USA, pp. 446-453

BEAUDOUIN-LAFON Michel. *Interaction Instrumentale : de la manipulation directe à la réalité augmentée*. IHM'97, 9èmes Journées sur l'Ingénierie de l'Interaction Homme-Machine. Cépadués Editions, 1997, pp. 97-104

BENKO Hrvoje, WILSON Andrew D. et **BAUDISCH Patrick.** *Precise selection techniques for multi-touch screens*. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (CHI '06), GRINTER Rebecca,

RODDEN Thomas, AOKI Paul, CUTRELL Ed, JEFFRIES Robin et OLSON Gary. ACM, New York, USA, 2006, pp. 1263-1272

BIER Eric A., STONE Maureen C., PIER Ken, BUXTON William et DEROSE Tony D. *Toolglass and magic lenses : the see-through interface*. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, (SIGGRAPH '93). ACM, New York, USA, 1993, pp. 73-80

BREWSTER Stephen, CHOCHAN Faraz et BROWN Lorna. *Tactile feedback for mobile interactions*. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (CHI '07). ACM, New York, USA, 2007, pp. 159-162

BUXTON Bill. *A touching story : A personal perspective on the history of touch interfaces past and future*. In *Society for Information Display (SID) Symposium Digest of Technical Papers* vol. 41, 2010, pp. 444-448

BUXTON Bill. *Multi-Touch Systems that I Have Known and Loved*. Microsoft Research, 2007, 16p. disponible ici : http://www.ekatetra.com/downloads/Chorded_History_BuxtonMultiTouch.pdf

BUXTON Bill. *The Long Nose of Innovation*. In *Businessweek*, Janvier 2008

BUXTON William. *Lexical and pragmatic considerations of input structures*. In *SIGGRAPH Computer Graphics*. 17, 1 1983, pp. 31-37

BUXTON William et MYERS Brad A. Study in two-handed input. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (CHI '86), Marilyn Mantei and Peter Orbeton. ACM, New York, NY, USA, 321-326. 1986

BUXTON William. *There's more to interaction than meets the eye : Some issues in manual input*. In *User Centered System Design : New Perspectives on Human-Computer Interaction*. Lawrence Erlbaum, Hillsdale, 1986, pp. 319-337

CADOZ Claude. *Le Geste canal de communication homme-machine : la communication instrumentale*. In *TSI (Traitement du Signal et des Images)*, vol. 13. 1994, pp. 31-61 (coll. Hermès)

CARD Stuart K., MACKINLAY Jock D. et ROBERTSON George G. *A morphological analysis of the design space of input devices*. ACM, 1991, pp. 99-122

CARD Stuart K., ENGLISH W. K. et BURR B. J. *Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys, for text selection on a CRT*. In *Human-computer interaction*, BAECKER R. M. Morgan Kaufmann Publishers Inc., San Francisco, USA, 1987, pp. 386-392

CHANG Angela et O'SULLIVAN Conor. Audio-haptic feedback in mobile phones. In CHI '05 extended abstracts on Human factors in computing systems (CHI EA '05). ACM, New York, USA, 2005, pp. 1264-1267

CLARKE Arthur C. *Extra-Terrestrial Relays published*. In *issue of Wireless World*, novembre 1945

COUTRIX Céline et NIGAY Laurence. *Mixed reality : a model of mixed interaction*. In *Proceedings of the working conference on Advanced visual interfaces* (AVI '06). ACM, New York, USA, pp. 43-50

DIETZ Paul et LEIGH Darren. *DiamondTouch : a multi-user touch technology*. In *Proceedings of the 14th annual ACM symposium on User interface software and technology* (UIST '01). ACM, New York, USA, 2001, pp. 219-226

EKMAN Paul et FRIESEN Wallace V. *The repertoire of nonverbal behavior : Categories, origins, usage, and coding*. In *Semiotica* (janvier 1969), pp. 49- 98

FITZMAURICE George W., ISHII Hiroshi et BUXTON William A. S. *Bricks : laying the foundations for graspable user interfaces*. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (CHI '95). ACM Press, New York, USA, pp. 442-449

FOLEY James D., WALLACE Victor L. et CHAN Peggy. *The human factors of computer graphics interaction techniques.* In *IEEE Computer Graphics and Applications*, 1984, pp. 13-48

GEORGE Ron et BLAKE Joshua. *Objects, Containers, Gestures, and Manipulations : universal foundational metaphors of natural user interfaces.* In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '10)*, Atlanta, Georgia, USA, 2010

GIBSON J. James. *Observations on active touch.* In *Psychological Review* (69), 1962, pp. 477-490

GIBSON James J. *The Theory of Affordances.* In *Perceiving, Acting, and Knowing.* Robert Shaw and John Bransford, New York, 1977, pp. 67-82

GUIARD Y. et FERRAND T. *Asymmetry in bimanual skills.* In *Manual asymmetries in motor performance*, ELLIOTT D. et ROY E.A. CRC Press, 1995, pp. 175-195

HAN Jefferson Y. *Low-cost multi-touch sensing through frustrated total internal reflection.* In *Proceedings of the 18th annual ACM symposium on User interface software and technology (UIST '05)*. ACM, New York, USA, 2005, pp. 115-118

HARDENBERG Christian von et BÉRARD François. *Bare-hand human-computer interaction.* In *Proceedings of the 2001 workshop on Perceptive user interfaces (PUI '01)*. ACM, New York, USA, 2001, pp. 1-8

HARLING Philip A. et EDWARDS Alistair D. N. *Hand Tension as a Gesture Segmentation Cue.* In *Proceedings of gesture workshop on progress in gestural interaction*, Springer-Verlag, London, UK, 1996, pp. 75-88

HUTCHINS Edwin L., HOLLAN James D., et NORMAN Donald A. *Direct manipulation interfaces.* In *Human-Computer Interaction* (1, 4), décembre 1985, pp. 311-338

ISHII Hiroshi. *Tangible bits : beyond pixels.* In *Proceedings of the 2nd international conference on Tangible and embedded interaction (TEI '08)*. ACM, New York, USA, 2008, chap. XV-XXV

ISHII Hiroshi et ULLMER Brygg. *Tangible bits : towards seamless interfaces between people, bits and atoms.* In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '97)*. ACM, New York, USA, 1997, pp. 234-241

ISHII Hiroshi et ULLMER Brygg. *The metaDESK : models and prototypes for tangible user interfaces.* In *Proceedings of the 10th annual ACM symposium on User interface software and technology (UIST '97)*. ACM, New York, USA, pp. 223-232

IZADI Shahram, HODGES Steve, BUTLER Alex, RRUSTEMI Alban et BUXTON Bill. *ThinSight : Integrated Optical Multi-touch Sensing through Thin Form-factor Displays.* In *Proceedings of the 2007 workshop on Emerging displays technologies : images and beyond : the future of displays and interaction (EDT '07)*. ACM, New York, USA, Art. 6

JACOB Robert J.K., GIROUARD Audrey, HIRSHFIELD Leanne M., HORN Michael S., SHAER Orit, SOLOVEY Erin Treacy, et ZIGELBAUM Jamie. *Reality-based interaction : a environnement de développement for post-WIMP interfaces.* In *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems (CHI '08)*. ACM, New York, USA, 2008, pp. 201-210

JACOB Robert J. K., SIBERT Linda E., McFARLANE Daniel C. et MULLEN M. Preston Jr. *Integrity and separability of input devices.* In *ACM Trans. Computer Human Interaction* (1, 1) , 1994, pp. 3-26

KALTENBRUNNER Martin, GEIGER Günter et JORDÀ Sergi. *Dynamic patches for live musical performance.* In *Proceedings of the 2004 conference on new interfaces for musical expression (NIME '04)*, Michael J. Lyons. National University of Singapore, Singapore, 2004, pp. 19-22

KALTENBRUNNER Martin, BOVERMANN T., BENCINA R., Costanza E. *TUIO - A Protocol for Table Based Tangible User Interfaces.* In *Proceedings of the 6th International Workshop on Gesture in Human-Computer Interaction and Simulation*, Vannes, France, 2005

KAMMER Dietrich, WOJDZIAK Jan, KECK Mandy, GROH Rainer et TARANKO Severin. *Towards a formalization of multi-touch gestures.* In *ACM International Conference on Interactive Tabletops and Surfaces (ITS '10)*. ACM, New York, USA, 2010, pp. 49-58

KARAM Maria et Schraefel M.C. *A taxonomy of gestures in human computer interaction.* In *Transactions on Computer-Human Interactions*, TBD, 2005

KENDON Adam. *Current issues in the study of gesture.* In *Biological Foundations of Gesture Hillsdale*. New Jersey, Lawrence Erlbaum Associates, 1986, pp. 23-47

KETTEBEKOV Sanshzar. *Exploiting prosodic structuring of coverbal gesticulation.* In *Proceedings of the 6th international conference on Multimodal interfaces (ICMI '04)*. ACM, New York, USA, 2004, pp. 105-112

KAY Alan. *A personal computer for children of all ages.* In *Proceedings of the ACM national conference*, Boston Aug. Xerox Palo Alto Research Center, 1972

KURTENBACH Gordon, FITZMAURICE George, BAUDEL Thomas et BUXTON Bill. *The design of a GUI paradigm based on tablets, two-hands, and transparency.* In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '97)*. ACM, New York, USA, 1997, pp. 35-42

LEE SK, BUXTON William et SMITH K. C. *A multi-touch three dimensional touch-sensitive tablet.* In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '85)*. ACM, New York, USA, 1985, pp. 21-25

MACALUSO Emiliano et DRIVER Jon. *Multisensory spatial interactions : a window onto functional integration in the human brain.* In *Trends in Neurosciences (28)* : 2005, pp. 263-271

McAVINNEY Paul. *The Sensor Frame - a gesture-based device for the manipulation of graphic objects.* In *Sensor Frame*, Carnegie-Mellon University, Pittsburgh, 1986

McKENZIE C. L. et IBERALL T. *The Grasping hand.* In *Advances in Psychology (104)*. 1994, 500 p.

MALIK Shahzad, RANJAN Abhishek et BALAKRISHNAN Ravin. *Interacting with large displays from a distance with vision-tracked multi-finger gestural input.* In *Proceedings of the 18th annual ACM symposium on User interface software and technology (UIST '05)*. ACM, New York, USA, 2005, pp. 43-52

MILGRAM Paul et KISHINO Fumio. *A taxonomy of mixed reality visual displays.* In *IEICE Transactions on Information Systems (Vol. E77-D, No. 12)*, 1994, 15p.

MINSKY Margaret, MING Ouh-young, STEELE Oliver, BROOKS Frederick P., Jr. et BEHENSKY Max. *Feeling and seeing: issues in force display.* In *Proceedings of the 1990 symposium on Interactive 3D graphics (I3D '90)*. ACM, New York, USA, 1990, pp. 235-241

NAKATANI Lloyd H. et ROHRLICH John A. *Soft machines : A philosophy of user-computer interface design.* In *Proceedings of the SIGCHI conference on Human Factors in Computing System (CHI '83)*. ACM, New York, USA, 1983, p. 19-23

NAPIER John R. *The prehensile movements of the human hand.* In *Surger* 38-B(4), 902-913. 1956. Disponible ici : <http://www.ncbi.nlm.nih.gov/pubmed/13376678>

NESPOULOUS J-L et LECOURES A. R. *Gestures : Nature and Function.* In *The Biological foundations of Gestures : Motor and Semiotic Aspects..*). Lawrence Erlbaum Associates, Hillsdale, 1986, pp. 49-62

NEWELL Allen, PERLIS A.J. et SIMON H.A. *What is computer science?* [Lettre à l'éditeur]. In *Science* 157, 1967, pp. 1373-1374

NORMAN Donald Arthur. *Affordances, conventions, and design.* In *Interactions*, Volume 6 Issue 3, 1999, pp. 38-42

NORMAN Donald Arthur. *Complexity versus Difficulty : Where should the Intelligence Be?*. Plenary Address. In *Proceedings of the 7th international conference on Intelligent user interfaces (IUI '02)*. ACM, New York, USA, 2002, pp. 4-4

O'REGAN J. Kevin et **NOË Alva.** *A sensorimotor account of vision and visual consciousness*. In *Behavioral and brain sciences*, 2001, pp. 939-1031

OWEN Russell, KURTENBACH Gordon, FITZMAURICE George, BAUDEL Thomas et **BUXTON Bill.** *When it gets more difficult, use both hands: exploring bimanual curve manipulation*. In *Proceedings of Graphics Interface (GI '05)*. Canadian Human-Computer Communications Society, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 2005, pp. 17-24

QUEK Francis, McNEILL David, BRYLL Robert, DUNCAN Susan, MA Xin-Feng, KIRBAS Cemil, McCULLOUGH Karl E. et **ANSARI Rashid.** *Multimodal human discourse : gesture and speech*. ACM, Computer-Human Interact (9, 3), 2002, pp. 171-193

QUEK Francis K. H. *A purpose taxonomy of gestural interaction*. Vision interfaces and systems laboratory, center for human compute interaction, non publié, 2003. Disponible ici : <http://vislab.cs.vt.edu/~quek/Classes/Aware+EmbodiedInteraction/PAPERS/Que03.pdf>

QUEK Francis K. H. *Toward a vision-based hand gesture interface*. In *Proceedings of the conference on Virtual reality software and technology (VRST '94)*. World Scientific Publishing Co., USA, 1994, pp. 17-31

QUEK Francis K. H. *Virtual reality software & technology : proceedings of the VRST '94*

RÉVÉSZ Géza. *System der optischen und haptischen Raumtäuschungen*. In *Zeitschrift für Physiologie*, 131, 1934, pp. 296-375

RIVIERE Jean-Baptiste (de la), KERVEGANT Cédric, DITTLO Nicolas, COURTOIS Mathieu et **ORVAIN Emmanuel.** *3D multitouch: when tactile tables meet immersive visualization technologies*. In *ACM SIGGRAPH 2010 Emerging Technologies (SIGGRAPH '10)*. ACM, New York, USA, 2010, 1 p.

SHNEIDERMAN Ben. *Direct Manipulation : a step beyond programming languages*. In *Computer* (16), 1983, pp. 57-69

SHNEIDERMAN Ben. *Software Psychology : human factors in computer and information systems*. In *Winthrop Computer Systems Series*. Winthrop Publishers, 1980

SALISBURY J. Kenneth et **SRINIVASAN Mandayam A.** *Phantom-Based Haptic Interaction with Virtual Objects*. In *IEEE Computer Graphics Applications*. 17, 1997, pp. 6-10

SASAKI L., Fedorkow, G., BUXTON William, RETTARETH C., et **Smith K.C.** *A Touch-Sensitive Input Device*. In *Proceedings of the Fifth International Conference on Computer Music*. North Texas State University, Texas, 1981

SEARS Andrew et **SHNEIDERMAN Ben.** *High precision touchscreens: design strategies and comparisons with a mouse*. In *International journal of man-machine studies*, 1991, pp. 593-613

SMITH David Canfield, IRBY Charles, KIMBALL Ralph et **HARSLEM Eric.** *The star user interface : an overview*. In *AFIPS Conference Proceedings*, vol. 55. National Computer Conference, AFIPS Press, Arlington, VA, USA, 1986, pp. 383-396

STURMAN David J. et **ZELTZER David.** 1993. A design method for "whole-hand" human-computer interaction. *ACM Trans. Inf. Syst.* (11, 3), 1993, pp. 219-238

SUTHERLAND Ivan Edward. *SketchPad : A man-machine graphical communication System*. In *AFIPS Spring Joint Computer Conference*. 1963, pp. 329-346

WEISER Mark. *The computer for the 21st century.* In *Human-computer interaction*, Ronald M. Baecker, Jonathan Grudin, William A. S. Buxton, and Saul Greenberg. San Francisco, USA, Weiser, pp. 933-940

Wobbrock Jacob O., Morris Meredith Ringel et Wilson Andrew D. *User-defined gestures for surface computing.* In *Proceedings of the 27th international conference on Human factors in computing systems (CHI '09)*. ACM, New York, USA, 2009, pp. 1083-1092

WEXELBLAT Alan. *Natural gesture in virtual environments.* In *Proceedings of the conference on Virtual reality software and technology (VRST '94)*. World Scientific Publishing Co., USA, 1994, pp. 5-16

WILLIAMS Gregg. *The Lisa Computer System.* In *Byte Magazine*. 1983, pp. 33-50

WILLIAMS Gregg. *The Apple Macintosh Computer.* In *Byte Magazine*. 1984, pp. 30-54

WILSON Andrew D. *TouchLight : an imaging touch screen and display for gesture-based interaction.* In *Proceedings of the 6th international conference on Multimodal interfaces (ICMI '04)*. ACM, New York, USA, 2004, pp. 69-76

ZIMMERMAN Thomas G., LANIER Jaron, BLANCHARD Chuck, BRYSON Steve et HARVILL Young. *A hand gesture interface device.* SIGCHI Bull. 17, SI, 1986, pp. 189-192

Thèses - DEA

ANSARI Anees. *Direct 3D Interaction Using A 2D Locator Device.* Thèse sous la direction de LES Piegł. Department of Computer Science And Engineering, College of Engineering, University of South Florida, 2003

FITZMAURICE George W. *Graspable User Interfaces.* Thèse, University of Toronto, Toronto, Canada, Canada. Sous la direction de William Buxton, 1996

MEHTA Nimish. *A Flexible Machine Interface.* Thèse, Department of Electrical Engineering, University of Toronto, 1982

MULLER Laurence Y. L. *Multi-touch displays : design, applications and performance evaluation.* Thèse, sous la direction de BELLEMAN R.G. Universiteit van Amsterdam, Amsterdam, 2008

RUBINE Dean Harris. *The Automatic Recognition of Gestures.* Thèse. Carnegie Mellon University, Pittsburgh, USA 1991

SMITH D.C. *Pygmalion : A computer program to model and stimulate creative thought.* Thèse. Basel, Stuttgart : Birkhauser Verlag, Stanford University Computer Science Department, 1975

STANLEY Douglas Edric. *Mode d'emploi : lexique de l'interactivité.* DEA Esthétique, Sciences et Technologies des Arts. Paris, Université de Paris VIII, 1997

Rapports

1394 Trade Association. DC 1394-based Digital Camera Specification Ver.1.31. 2004. Disponible ici : http://damien.douxchamps.net/ieee1394/libdc1394/iidc/IIDC_1.31.pdf

ACM SIGCHI. *Curricula for Human-Computer Interaction.* Rapport technique. ACM, New York, USA, 1992, disponible ici : http://old.sigchi.org/cdg/cdg2.html#2_1

COLEGROVE Jennifer. *The State of the Touch-Screen Market in 2010.* In *Display marketplace*. Mars 2010, pp. 22-24

COLEGROVE Jennifer, HSIEH Calvin. *Touch Panel Market Analysis Report.* DisplaySearch, 2010

DENNING Peter, COMER Douglas E., GRIES David, MULDER Michael C., TUCKER Allen B., TURNER A. Joe et YOUNG Paul R. *Computing as a discipline : preliminary report of the ACM task force on the core of computer science.* In *SIGCSE Bulletin* (20, 1), 1998, pp. 41-41

ENGELBART Douglas C. *Augmenting human intellect : a conceptual environnement de développement.* Report. Air force office of scientific research, Washington 25, C.C., 1962, 133 p.

National Research Council. *Innovation in Information Technology.* In *Computer Science and Telecommunications Board*. Washington DC, The National Academies Press, 2003, 84 p.

NIELSEN Jakob. *CHI'88 Trip Report.* Washington D.C., 15-19 mai 1988

SCHÖNING Johannes et al. In *Technical Report TUM-I0833, Technical Reports of the Technical, University of Munich*, 2008

Documentaires - Interviews vidéo

STIEGLER Bernard. *The Ister.* Documentaire de David Barison and Daniel Ross. 2004, 189 min.

GIBSON William. Interview, CNNfn, 26 août 1997

BUXTON Bill. *CES 2010 : NUI with Bill Buxton.* Larry Larsen et Bill Buxton. Channel9. Disponible sur : <http://channel9.msdn.com/Blogs/LarryLarsen/CES-2010-NUI-with-Bill-Buxton>

Encyclopédies - Dictionnaires

Dictionnaire Historique de la Langue Française. Dictionnaire historique. Sous la direction d'Alain Rey. Le Robert, 2006, 4304 p.

Le Petit Robert. *Dictionnaire alphabétique et analogique de la langue française.* Le Robert, (Coll. Dictionnaires Generalistes), 1996

Ressources en ligne

CASSINELLI Alvaro. *The khronos projector. A video time-warping machine with a tangible deformable screen.* Disponible ici : <http://www.k2.t.u-tokyo.ac.jp/members/alvaro/Khronos/>

NUI Group. Disponible ici : <http://www.nuigroup.com/>

MULTIGESTURE.NET. Disponible ici : <http://www.multigesture.net>

REACTABLE. Disponible ici : <http://www.reactable.com>

BOX2D. Disponible ici : <http://www.box2d.org/>

TABLE DES ILLUSTRATIONS

Figure 1 - Interface graphique du Xerox Star reprenant les principes déjà présents dans l'ordinateur Alto, 1981	38
Figure 2 - Le gant de données à fibres optiques conçu par Thomas Zimmerman en 1987, VPL Research.....	49
Figure 3 - Bras haptique « Phantom » de la société SensAble	50
Figure 4 - Représentation du concept d'interface palpable. Reproduction, [Ishii <i>et al.</i> , 95].....	52
Figure 5 - Représentation simplifiée du continuum de virtualité. Reproduction, [Milgram 94]	53
Figure 6 - Dispositif metaDESK. In [Ishii et Ullmer, 1997].....	57
Figure 7 - Incarnation physique des éléments du GUI dans les TUIs. In [Ishii et Ullmer, 97]	58
Figure 8 - ReactTable : instrument tangible interactif. Jordà <i>et al.</i> , 2005.....	59
Figure 9 - Interacteur tangible mixte utilisé sur la Reactable	60
Figure 10 - La Reactable en cours d'utilisation.....	61
Figure 11 - Evolution simplifiée des IHM.....	66
Figure 12 - Les phases du processus d'innovation en fonction du temps	69
Figure 13 - Terminal PLATO IV, 1972	70
Figure 14 - Videodesk de Myron Krueger, 1983	72
Figure 15 - Tablette multi-tactile capacitive. Bell Labs, 1985	73
Figure 16 - SensorFrame.....	74
Figure 17 - VideoDesk.....	75
Figure 18 - La tablette iGesture Pad. Le clavier TouchStream. Fingerworks	75
Figure 19 - La table multi-tactile DiamonTouch développée par les laboratoires de Mitsubishi Electric Research, 2001.....	76
Figure 20 - A gauche, le contrôleur Lemur. A droite le Tenori-on	77
Figure 21 - TouchLight.....	78
Figure 22 - Microsoft Surface V1	79
Figure 23 - Prototype T1, 2003. Microsoft Surface V1, 2007	80
Figure 24 - CCV. Traitement d'image.....	81
Figure 25 - Jeff Han au TED. « MagicWall » de Perceptive Pixel.....	81
Figure 26 - Dynabook d'Alan Kay. Newton MessagePad d'Apple.....	82
Figure 27 - Prototype « Tab », 1991. iPhone 4, 2010	84
Figure 28 - Prototype « Pad », 1991. Tablette iPad d'Apple, 2010.....	85
Figure 29 - Prévisionnel des revenus générés par les écrans tactiles.....	87
Figure 30 - Technologie ThinSight	88
Figure 31 - Table interactive SUR40 ou Microsoft Surface V2	89
Figure 32 - Videoplace, Myron Kruger, 1974 et Figure 33 - Khronos Projector, Alvaro Cassinelli, 2005.....	90
Figure 34 - Exemple de programmation graphique réalisé avec EyesWeb 5.0.1.4	92
Figure 35 - Tron, 1982. The Island, 2005	95
Figure 36 - Minority Report, Steven Spielberg, 2002.....	95
Figure 37 - Evolution du nombre de recherches du terme multi-touch.....	98
Figure 38 - Dispositif de captation gestuelle	99
Figure 39 - Classification des interactions homme-machine gestuelle. Maria Karam et M.C. Schraefel.....	103
Figure 40 - Illustration simplifiée du fonctionnement d'un écran capacitif	106
Figure 41 - Schéma des matrices conductrices d'un écran résistif	107
Figure 42 - Configuration d'un système FTIR.....	110

Figure 43 - Configuration d'un système DI	111
Figure 44 - Configuration d'un système DSI	112
Figure 45 - Configuration d'un système LLP	113
Figure 46 - Illustration simplifiée du fonctionnement d'un cadre optique infrarouge.....	114
Figure 47 - Homoncule sensitif.....	122
Figure 48 - Sculpture de l'homoncule moteur.....	122
Figure 49 - Représentation du continuum de Kendon. Reproduction. In [McNEILL, 1996].....	124
Figure 50 - Taxinomie générale du geste	134
Figure 51 - Classification proposée par Ron George et Joshua Blake	140
Figure 52 - Représentation du couplage de la phrase gestuelle avec les états du système	148
Figure 53 - Articulations de la main.....	152
Figure 54 - Mouvements articulaires de la main. Schéma de l'American Society for Surgery of Hand, 1978.....	153
Figure 55 - Éventail de dispositifs multi-tactiles. De la plus petite surface à la plus grande	162
Figure 56 - Exemple de geste atomique	166
Figure 57 - Illustration du geste d'écriture	169
Figure 58 - Illustration du geste de dessin. Manuel et bi-manuel	170
Figure 59 - Illustration de quelques gestes de manipulation rappelant le style WIMP	171
Figure 60 - Illustration de quelques gestes de manipulation spécifiques.....	172
Figure 61 - Illustration de quelques gestes lexicaux.....	173
Figure 62 - Vue avant du prototype T1 multi-tactile	189
Figure 63 - Illuminateur infrarouge conçu en interne	193
Figure 64 - Configuration du système de capture optique.....	194
Figure 65 - Sensibilité spectrale du système optique utilisé pour le prototype	196
Figure 66 - Arbre des choix technologiques	197
Figure 67 - Captures d'images du logiciel CCV	198
Figure 68 - Capture d'image du logiciel CCV. Chaîne de filtre appliquée à l'image brute issue de la caméra	199
Figure 69 - Capture d'écran de l'interface de <i>TuioSimulator</i>	203
Figure 70 - Illustrations de la vue caméra	204
Figure 71 - Grille de calibration projetée sur la surface déformée.....	204
Figure 72 - Illustration de la correction barycentrique.....	205
Figure 73 - Représentation des mesures de latence effectuée sur le vidéoprojecteur et l'écran	207
Figure 74 - Table interactive. Borne multi-tactile.....	212
Figure 75 - <i>TouchDataListener</i> , outil de lecture d'enregistrements tactiles	215
Figure 76 - Types d'interaction utilisés dans le projet	217
Figure 77 - Capture d'écran du projet <i>Andromède</i>	217
Figure 78 - Types d'interaction utilisés dans le projet	218
Figure 79 - Capture d'écran du projet <i>TouchMap</i>	218
Figure 80 - Types d'interaction utilisés dans le projet	219
Figure 81 - Capture d'écran du projet <i>SPhIre</i>	219
Figure 82 - Types d'interaction utilisés dans le projet	220
Figure 83 - Capture d'écran de deux projets avec un menu de type carrousel.....	220
Figure 84 - Types d'interaction utilisés dans le projet	221
Figure 85 - Capture d'écran du projet LUNAR	221
Figure 86 - Types d'interaction utilisés dans le projet	222

Figure 87 - Capture d'écran du projet GSMSS (trois vues utilisateur)	222
Figure 88 - Types d'interaction utilisés dans le projet	223
Figure 89 - Capture d'écran du projet GSMSS version table	223
Figure 90 - Différence d'échelle observée entre l'interaction et l'objet d'intérêt.....	224
Figure 91 - Observation de points aberrants.....	227
Figure 92 - Illustration du changement de l'ID des points de contact au cours de la phrase gestuelle.....	228
Figure 93 - bloc-diagramme des architectures logicielles multi-tactiles existantes	231
Figure 94 - Organigramme fonctionnel des architectures logicielles multi-tactiles existantes	232
Figure 95 - Organigramme simplifié de l'architecture de l'intergiciel	233
Figure 96 - Organigramme fonctionnel de l'environnement de développement	236
Figure 97 - Diagramme de classe de l'intergiciel	237
Figure 98 - Organigramme fonctionnel d'extraction et de conversion des données tactiles.....	238
Figure 99 - Illustration de la synchronisation entre l'intergiciel et l'application	244
Figure 100 - Cycle de gestion des événements tactiles	247
Figure 101 - Illustration simplifiée de l'algorithme de correction des points	257
Figure 102 - Panneau de modification des paramètres du SorClient sous <i>Unity</i> 3D	262
Figure 103 - Enregistrement des patrons d'interaction sous <i>Unity</i>	267
Figure 104 - Inscription 3D de la main dans l'espace virtuel	282
Figure 105 - Cycles de l'innovation concernant le domaine des TIC	286
Figure 106 - <i>VideoHarp</i> , produite par la société <i>SensorFrame</i>	287
Figure 107 - Concept art d'avant-projet. Croquis réalisé par Jean-François Pal	290
Figure 108 - Scène 3D issue du concept art.....	290
Figure 109 - Concept art d'un prototype d'écrans multi-tactiles circulaires translucides.....	291
Figure 110 - Concept art d'un prototype de réalité augmentée.....	291
Figure 111 - Extrait de la plaquette de présentation du projet LUNAR.....	292
Figure 112 - Extrait de la plaquette de présentation du projet AQUARIUS.....	293
Figure 113 - Photo du showroom réalisé pour DCNS, site de Toulon	294
Figure 114 - Photo du showroom réalisé pour DCNS, site de Toulon	294
Figure 115 - Vue d'ensemble de la démonstration réalisée devant le pôle communication et R&D de DCNS Toulon	296
Figure 116 - Vue d'ensemble de la démonstration réalisée devant le pôle communication et R&D de DCNS Toulon	296
Figure 117 - Vue d'ensemble du showroom DCNS au salon Milipol 2009.....	297
Figure 118 - Simulation 3D multi-tactile présentée lors du salon Milipol 2009	297
Figure 119 - Un père apprenant à son fils comment utiliser l'interface du projet LUNAR	298
Figure 120 - Un utilisateur en train de jouer au jeu Warcraft III adapté pour le multi-tactile.....	298
Figure 121 - Simulation 3D multi-tactile présentée lors du salon ARICAD	299
Figure 122 - Conférence réalisée lors du salon ARICAD	299
Figure 123 - Photo prise durant le salon Laval Virtual.....	300
Figure 124 - Photo prise durant le salon Laval Virtual.....	300

Liste des tableaux

Tableau 1 - Domaines d'application des technologies multi-tactiles et répartition des revenus par secteur	104
Tableau 2 - Élément de langage de la formalisation de Kammer <i>et al.</i> In [KAMMER, 2010]	166

TABLE DES MATIÈRES

REMERCIEMENTS	3
AVANT-PROPOS	7
ORIGINE DE LA RECHERCHE	7
CONTINUE DE MA RECHERCHE	9
CADRE PROFESSIONNEL	10
LA THESE, JUSTE UN COMMENCEMENT.....	10
SOMMAIRE	12
INTRODUCTION.....	14
LE SYSTEME TECHNICIEN	15
DESIGN CENTRE UTILISATEUR	16
ORIGINE DE LA RECHERCHE.....	17
ÉTAT DE LA RECHERCHE	19
METHODOLOGIE DE LA RECHERCHE	21
ANNONCE DU PLAN.....	22
CHAPITRE I CHAMP DE RECHERCHE.....	26
TERMINOLOGIE GENERALE.....	29
1. INTERACTION HOMME-MACHINE	31
1.1. INTERACTION HOMME-MACHINE.....	31
1.2. INTERFACE WIMP	33
1.3. LIMITES DU MODELE WIMP	38
2. INSTRUMENT D'INTERACTION	40
2.1. PERIPHERIQUE D'ENTREE	40
2.2. INTERACTION INSTRUMENTALE.....	42
2.3. INSTRUMENT HAPTIQUE	45
<i>Terminologie.....</i>	<i>46</i>
<i>Réalisme et simulation physique</i>	<i>46</i>
<i>Exemple 1 : le gant de données</i>	<i>48</i>
<i>Exemple 2 : le bras haptique.....</i>	<i>49</i>
<i>Limitations des technologies haptiques.....</i>	<i>51</i>
2.4. INSTRUMENT MIXTE.....	51
<i>Les « bits tangibles ».....</i>	<i>53</i>
<i>Terminologie.....</i>	<i>54</i>

<i>Caractérisation des interfaces tangibles</i>	56
<i>Exemple 1 : MetaDesk, 1997</i>	57
<i>Exemple 2 : ReacTable - 2005</i>	58
<i>Limitation des TUI</i>	62
CHAPITRE II INTERFACES TACTILES MULTI-POINTS	64
INTRODUCTION	65
INTERFACES UTILISATEUR NATURELLES (NUI)	65
INFORMATIQUE UBIQUITAIRE	66
1. UNE LONGUE PERIODE D'INCUBATION	68
2. HISTOIRE DES TECHNOLOGIES MULTI-TACTILES	70
<i>1960 – 1991 : une phase expérimentale</i>	70
<i>1998 – 2003 : Premières commercialisations grand public</i>	75
<i>2003 – 2007 : laboratoire Microsoft</i>	77
<i>2005 – 2010 : Jeff Han et Apple</i>	80
<i>Retour en 1991 : Mark Weiser</i>	83
<i>Aujourd'hui</i>	86
3. PRE-APPROPRIATION TECHNOLOGIQUE	89
3.1. <i>L'art numérique</i>	90
3.2. <i>Les univers de science-fiction</i>	93
3.3. <i>Les communautés de chercheurs</i>	97
3.4. <i>Le monde vidéo-ludique</i>	99
CHAPITRE III ESPACE DE CONCEPTION	101
1. DOMAINES D'APPLICATION	103
2. PERIPHERIQUES D'ENTREE MULTI-TACTILES	104
2.1. TECHNOLOGIES NON PERCEPTIVES	105
<i>Technologie capacitive</i>	105
<i>Technologie résistive</i>	106
2.2. TECHNOLOGIES PERCEPTIVES	107
<i>Spectre lumineux</i>	108
<i>Réflexion totale frustrée</i>	109
<i>Illumination diffuse</i>	110
<i>Illumination de surface diffuse (DSI)</i>	112
<i>Plan de lumière laser infrarouge</i>	112
2.3. TECHNOLOGIES A CAPTEURS OPTIQUES	113

3.	PERIPHERIQUES DE SORTIE	115
3.1.	CANAL VISUEL	115
3.2.	CANAL AUDIO.....	116
3.3.	CONFIGURATION.....	116
4.	LE CANAL GESTUEL	117
4.1.	DEFINITION.....	118
4.2.	INTEGRATION SENSORI-MOTRICE	120
4.3.	CONTINUUM DE CONVENTIONNALITE	123
4.4.	TAXINOMIE FONCTIONNELLE DU GESTE	125
4.4.1.	<i>Fonction épistémique</i>	126
4.4.2.	<i>Fonction ergotique</i>	128
4.4.3.	<i>Fonction sémiotique</i>	130
CHAPITRE IV	MODÈLE D'INTERACTION	135
1.	INTERFACE UTILISATEUR NATURELLE	138
1.1.	TERMINOLOGIE.....	138
1.2.	METAPHORE DU REEL	141
1.3.	LIMITATIONS DE LA METAPHORE DU REEL	144
2.	MODELISATIONS TEMPORELLES DU GESTE.....	146
2.1.	FORME CANONIQUE DU GESTE	146
2.2.	ETATS DU SYSTEME.....	147
2.3.	DISCRET / CONTINU	149
3.	MODELISATION SPATIALE DU GESTE	149
3.1.	INTERFACE MULTIPLEXEE DANS L'ESPACE.....	149
3.2.	DEGRES DE LIBERTE	151
3.3.	LE DEGRE D'INDIRECTION	154
3.4.	LE DEGRE D'INTEGRATION	154
3.5.	LE DEGRE DE COMPATIBILITE.....	156
4.	MODALITES DES INTERACTIONS.....	157
4.1.	INTERACTION DIRECTE	157
4.2.	MORPHOLOGIE ET USAGE	157
4.3.	DEFINITION DE DONNEES	163
5.	TAXINOMIE DU GESTE	165
5.1.	FORMALISATION SYNTAXIQUE	165
5.2.	TAXINOMIE DES INTERACTIONS MULTI-TACTILES	167

6.	DESIGN D'INTERACTION	173
6.1.	VISIBILITE	174
6.2.	MAPPING NATUREL	176
6.3.	AFFORDANCE	177
6.4.	FEEDBACK.....	179
6.5.	GESTION DES ERREURS	183
CHAPITRE V	PROTOTYPES	186
1.	PROTOTYPAGE	187
1.1.	PROTOTYPE D'ECRAN LARGE MULTI-TACTILE	188
1.2.	TECHNIQUE D'ILLUMINATION.....	189
1.3.	SYSTEME DE CAPTURE OPTIQUE.....	193
2.	PREREQUIS LOGICIELS	198
2.1.	VISION PAR ORDINATEUR.....	198
2.2.	DETECTION ET SUIVI DES POINTS.....	199
2.3.	PROTOCOLE TUIO	201
2.4.	TUIOSIMULATOR.....	202
3.	CONSIDERATIONS TECHNIQUES	203
3.1.	DEFORMATION OPTIQUE	203
3.2.	LATENCE DU SYSTEME	206
	<i>Latence matérielle</i>	<i>206</i>
	<i>Latence logicielle</i>	<i>208</i>
	<i>Latence générale</i>	<i>208</i>
CHAPITRE VI	ENVIRONNEMENT DE DÉVELOPPEMENT.....	209
1.	EXTENSION DU MODELE D'INTERACTION	211
1.1.	METHODOLOGIE DE RECUEIL ET DE TRAITEMENT DES DONNEES.....	212
	<i>Matériel</i>	<i>212</i>
	<i>Logiciel.....</i>	<i>213</i>
1.2.	RECUEIL DES DONNEES	216
1.3.	AMBIGUÏTE D'ECHELLE	224
1.4.	POINTS ABERRANTS	226
1.5.	SENSIBILITE DES SYSTEMES OPTIQUES	228
2.	DÉVELOPPEMENT DE L'INTERGICIEL MULTI-TACTILE	231
2.1.	ARCHITECTURE LOGICIELLE	231
	<i>Limitations de l'existant.....</i>	<i>231</i>

<i>Proposition</i>	232
<i>Cadre de développement</i>	234
2.2. DONNEES TACTILES.....	238
<i>Récupération des données tactiles</i>	238
<i>CCV : Protocole TUIO</i>	238
<i>API PQLabs</i>	239
<i>API NextWindow</i>	240
<i>Intergiciel : la classe Device</i>	241
<i>Intergiciel : format des données tactiles</i>	242
2.3. ARCHITECTURE RESEAU	243
2.3.1. <i>Système de communication asynchrone</i>	243
2.3.2. <i>Paquet « TouchInfo »</i>	248
2.3.3. <i>Paquet de Patrons et de Paramètres d'interaction (PPP)</i>	249
2.3.4. <i>Patron et Paramètres d'Interaction (PPI)</i>	249
2.3.5. <i>Paquet « InteractionEvent »</i>	253
2.4. TRAITEMENT ET LISSAGE	255
2.4.1. <i>Correction des points et delta</i>	255
2.4.2. <i>Lissage de points</i>	258
3. IMPLEMENTATION DES API.....	259
3.1. <i>La classe SorClient</i>	260
3.2. <i>Interrogation de la couche graphique</i>	262
3.3. <i>Enregistrement des OCGM</i>	265
CONCLUSION.....	268
APPROCHE THEORIQUE	269
APPROCHE PRATIQUE	273
OUVERTURE THEORIQUE	276
TRAVAUX FUTURS	278
BILAN DU CIFRE	282
ANNEXES	284
1. ANNEXES GENERALES	285
1.1. CYCLES D'INNOVATION DE QUELQUES TECHNOLOGIES	286
1.2. VIDEOHARP	287
1.3. EXTRAIT COMPLET DE « 2001, L'ODYSEE DE L'ESPACE » D'ARTHUR C. CLARKE	288
2. DESIGN ET CONCEPTS ART	289
2.1. PROTOTYPE « INTOUCH »	290
2.2. PROTOTYPES DIVERS	291

2.3.	PROJET « <i>LUNAR</i>	292
2.4.	PROJET « <i>QUARIUS</i> »	293
2.5.	SALON DCNS - TOULON, LE MOURILLON	294
3.	SALONS ET DEMONSTRATIONS	295
3.1.	DCNS - DEMONSTRATION INTERNE – TOULON, 2009	296
3.2.	MILIPOL - PARIS, 2009.....	297
3.3.	TOULOUSE GAME SHOW – TOULOUSE, 2009	298
3.4.	ARICAD - PARIS, 2010.....	299
3.5.	LAVAL VIRTUAL - LAVAL, 2010.....	300
	BIBLIOGRAPHIE	301
	OUVRAGES GENERAUX	302
	REVUES SCIENTIFIQUES - ACTES DE CONFERENCE	303
	THÈSES - DEA	308
	RAPPORTS.....	308
	DOCUMENTAIRES - INTERVIEWS VIDEO	309
	ENCYCLOPÉDIES - DICTIONNAIRES.....	309
	RESSOURCES EN LIGNE.....	309
	TABLE DES ILLUSTRATIONS	310
	LISTE DES TABLEAUX	313
	TABLE DES MATIÈRES	314